

# DIGITAL MORPHOGENESIS VIA SCHELLING SEGREGATION

GEORGE BARMPALIAS, RICHARD ELWES, AND ANDY LEWIS-PYE

**ABSTRACT.** Schelling’s model of racial segregation has been extensively studied but has largely resisted rigorous analysis. According to the model, a large number of individuals of two types have their behaviour specified by two parameters. First, each is concerned only with their own ‘neighbourhood’, which is the set of individuals within a certain distance  $w$  of their location. Second, each has an ‘intolerance level’  $\tau$ , which is the proportion of their neighbourhood which they require to be their own type in order to be happy. Unhappy pairs of individuals are then chosen uniformly at random and are given the opportunity to swap locations, causing large segregated regions to form from an initially mixed configuration.

In [2], Brandt, Immorlica, Kamath and Kleinberg provided the first rigorous analysis of the 1 dimensional unperturbed version of Schelling’s model, for the case  $\tau = 0.5$ . Here we provide a rigorous analysis of the model’s behaviour more generally for  $\tau \in [0, 1]$ , and establish that as one varies  $\tau$ , surprising forms of threshold behaviour result, notably the existence of situations where an *increased* level of intolerance leads almost certainly to *decreased* segregation.

## 1. INTRODUCTION

A major achievement of the economist and game theorist Thomas Schelling was an elegant model of racial segregation, first described in 1969 [17]. This work is often cited (see for instance [7]) as one of the first agent-based computational economic models, and variants continue to be a popular topic among researchers in this discipline [7, 4, 8, 9]. It is also interesting to observe that Schelling’s model can be understood within the framework of *morphogenesis*, the analysis of how structure can arise from an initially random, or near random configuration. There has been work on similar ideas from other quarters. For instance, while Alan Turing is best known in the mathematical and computer science communities for his work formalising the algorithmically calculable functions, his most cited work [18] actually relates to morphogenesis. Turing was interested in understanding certain biological processes: the gastrulation phase of embryonic development, the process

---

Authors are listed alphabetically. Andy Lewis-Pye was previously called Andrew Lewis and was supported by a Royal Society University Research Fellowship. Barmpalias was supported by the Research Fund for International Young Scientists number 611501-10168 from the National Natural Science Foundation of China, and an International Young Scientist Fellowship number 2010-Y2GB03 from the Chinese Academy of Sciences; partial support was also received from the project Network Algorithms and Digital Information number ISCAS2010-01 from the Institute of Software, Chinese Academy of Sciences.

whereby dappling effects arise on animal coats, and phyllotaxy, i.e. the arrangement of leaves on plant stems.

Schelling's model looks to describe how individuals of different races come to organise themselves geographically within a community, into segregated regions, each of largely one race. Part of the significance of this model at the time was that it provided proof of a recurrent theme in Schelling's work – that the behaviour of individuals according to their interests at the local level, can produce results at the global level which are undesired by all. Running small simulations of his model, for example, Schelling observed that large levels of segregation can result in communities, when each individual has no preference for segregation, but still requires a certain low proportion of their own local neighbourhood to be of their own race. As noted by Zhang in [25] (which includes a good survey of research in the area), Schelling's model has every ingredient of a great theoretical work: it addresses an important real world issue, is simple to understand, produces unexpected results and it has deep implications for various social sciences. Although the explicit concern is racial segregation, the analysis is sufficiently abstract that any situation in which objects of two types arrange themselves geographically according to a certain preference not to be of a minority type in their neighbourhood, could constitute an interpretation. As pointed out in [20], for example, Schelling's model can be seen as a finite difference version of a differential equation describing interparticle forces. While the model has been extensively studied, see also for example [1, 13, 6, 11, 12, 16, 23, 24], it has been difficult to rigorously prove that the model produces the segregation behaviour clearly observed when running simulations.

We concentrate here on the one-dimensional version of the model, as in [2]. The model works as follows. One begins with a large number  $n$  of nodes (individuals) arranged in a circle. Each node is initially assigned a *type*, and has probability  $\frac{1}{2}$  of being of type  $\alpha$  and probability  $\frac{1}{2}$  of being of type  $\beta$  (the types of distinct individuals being independently distributed). We fix a parameter  $w$ , which specifies the 'neighbourhood' of each node in the following way: at each point in time the neighbourhood of the node  $u$ , denoted  $\mathcal{N}(u)$ , is the set containing  $u$  and the  $w$ -many closest neighbours on both sides – so the neighbourhood consists of  $2w + 1$  many nodes in total. The second parameter  $\tau \in [0, 1]$  specifies the proportion of a node's neighbourhood which must be of their type before they are happy. So, at any given moment in time, we define  $u$  to be *happy* if at least  $\tau(2w + 1)$  of the nodes in  $\mathcal{N}(u)$  are of the same type as  $u$ . One then considers a discrete time process, in which, at each stage, one pair of unhappy individuals of opposite types are selected uniformly at random and are given the opportunity to swap locations. We work according to the assumption that the swap will take place as long as each member of the pair has at least as many neighbours of the same type at their new location as at their former one (note that for  $\tau \leq \frac{1}{2}$  this will automatically be the case). The process ends when (and if) one reaches a stage at which there are no longer unhappy individuals of both types.

Much of the difficulty in providing a rigorous analysis stems from the lack of a well defined unique stable distribution for the underlying Markov process. Various authors have therefore worked with variants of the model in which perturbations are introduced into the dynamics so as to avoid this problem, i.e. the model is altered by introducing a further random element, allowing individuals to sometimes make moves which are detrimental with respect to their utility function. Such changes in the model might be justified by the assumption that we are dealing with individuals of ‘bounded rationality’ – one might consider that precise information concerning the racial composition of each neighbourhood is not available, for example. Young used techniques from evolutionary game theory – an analysis in terms of stochastically stable states – to develop the first results along these lines [22]. In a paper by Pancs and Vriend [16] and in a number of papers by Zhang [23, 24, 25], these ideas have been substantially developed, with the results generally being that complete global segregation is the only possible long run outcome.

In [2], Brandt, Immorlica, Kamath and Kleinberg used an analysis of locally defined stable configurations, combined with results of Wormald [21], to provide the first rigorous analysis of the unperturbed one-dimensional Schelling model, for the case  $\tau = \frac{1}{2}$ . The results obtained there are very different to those for the perturbed models: in the final configuration the average length of maximal segregated region is independent of  $n$  and only polynomial in  $w$ . So now the local dynamics do not induce global segregation in proportion to the size of the society but instead only a small degree of segregation at the local level. The suggestion is that these results are in accord with empirical studies of residential segregation in large populations [5, 10, 19].

In this paper we shall consider what happens more generally for  $\tau \in [0, 1]$  (for the unperturbed model), and we shall observe, in particular, that some remarkable threshold behaviour occurs. While some aspects of the approach from [2] remain, in particular the focus on locally defined stable configurations which can be used to understand the global picture, the specific methods of their proof (the use of ‘firewall incubators’, and so on) are entirely specific to the case  $\tau = \frac{1}{2}$ , and so largely speaking we shall require different techniques here. The picture which emerges is one in which one observes different behaviour in five regions. For  $\kappa$  which is the unique solution in  $[0, 1]$  to:

$$\left(\frac{1}{2} - \kappa\right)^{1-2\kappa} = (1 - \kappa)^{2-2\kappa},$$

(which is just slightly less than  $\sqrt{2}/4$ ) these regions are: (i)  $\tau < \kappa$ , (ii)  $\tau = \kappa$ , (iii)  $\kappa < \tau < \frac{1}{2}$ , (iv)  $\tau = \frac{1}{2}$ , (v)  $\tau > \frac{1}{2}$ . In fact we shall not consider the case  $\tau = \kappa$ , but the behaviour for all other values of  $\tau$  is given by the theorems below. Perhaps the most surprising fact is that, in some cases, increasing  $\tau$  almost certainly leads to decreased segregation. The assumption is always that we work with  $n \gg w$ , i.e. all results hold for all  $n$  which are sufficiently large compared to  $w$ . A *run* of length  $d$  is a set of  $d$ -many

consecutive nodes all of the same type. *Complete segregation* refers to any configuration in which there exists a single run to which all  $\alpha$  nodes belong.

**Theorem 1.1.** *Suppose  $\tau < \kappa$  and  $\epsilon > 0$ . For all sufficiently large  $w$ , if a node  $u$  is chosen uniformly at random, then the probability that any node in  $\mathcal{N}(u)$  is ever involved in a swap is  $< \epsilon$ . Thus there exists a constant  $d$  such that, for sufficiently large  $w$ , the probability  $u$  belongs to a run of length  $> d$  in the final configuration is  $< \epsilon$ .*

**Theorem 1.2.** *Suppose  $\tau \in (\kappa, \frac{1}{2})$  and  $\epsilon > 0$ . There exists a constant  $d$  such that (for all  $w$  and  $n \gg w$ ) the probability that  $u$  chosen uniformly at random will belong to a run of length  $\geq e^{w/d}$  in the final configuration, is greater than  $1 - \epsilon$ .*

**Theorem 1.3** ([2]). *Suppose  $\tau = \frac{1}{2}$ . There exists a constant  $c < 1$  such that for all  $\lambda > 0$ , the probability that  $u$  chosen uniformly at random will belong to a run of length greater than  $\lambda w^2$  in the final configuration, is bounded above by  $c^\lambda$ .*

Note that, if  $\frac{1}{2} < \tau \leq \frac{w+1}{2w+1}$ , then the process is identical to that for  $\tau = \frac{1}{2}$ , since in both cases a node requires  $w + 1$  many nodes of its own type in its neighbourhood in order to be happy.

**Theorem 1.4.** *Suppose that  $\tau > \frac{1}{2}$ , and that  $w$  is sufficiently large that  $\tau > \frac{w+1}{2w+1}$  (so that the process is not identical to that for  $\tau = \frac{1}{2}$ ). Then, with probability tending to 1 as  $n \rightarrow \infty$ , the initial configuration is such that complete segregation eventually occurs with probability 1.*

We have constructed a program which efficiently simulates the process. The outcomes of some simulations are illustrated in Figure 1. In the processes depicted here the number of nodes  $n = 50000$ , and in the diagrams individuals of type  $\alpha$  are coloured light grey and individuals of type  $\beta$  are coloured black. The inner ring displays the initial mixed configuration (in fact the configuration is sufficiently mixed that changes of type are not really visible, so that the inner ring appears dark grey). The outer ring displays the final configuration. Just immediately exterior to the innermost ring are second and third inner rings, which display individuals which are unhappy in the initial configuration and individuals belonging to ‘stable’ intervals in the initial configuration respectively (stable intervals will be defined subsequently, and in fact, this third ring is empty in these examples except for the case  $\tau = 0.3$ ). The process by which the final configuration is reached is indicated in the space between the inner rings and the outer ring in the following way: when an individual changes type this is indicated with a mark, at a distance from the inner rings which is proportional to the time at which the change of type takes place. In fact, for the case  $\tau > \frac{1}{2}$  one has to be a little careful in talking about the ‘final’ configuration – it is almost certainly the case that the numbers of each type in the initial configuration mean there will always be unhappy individuals of both types, but once a completely segregated configuration is reached all future configurations must remain completely segregated.

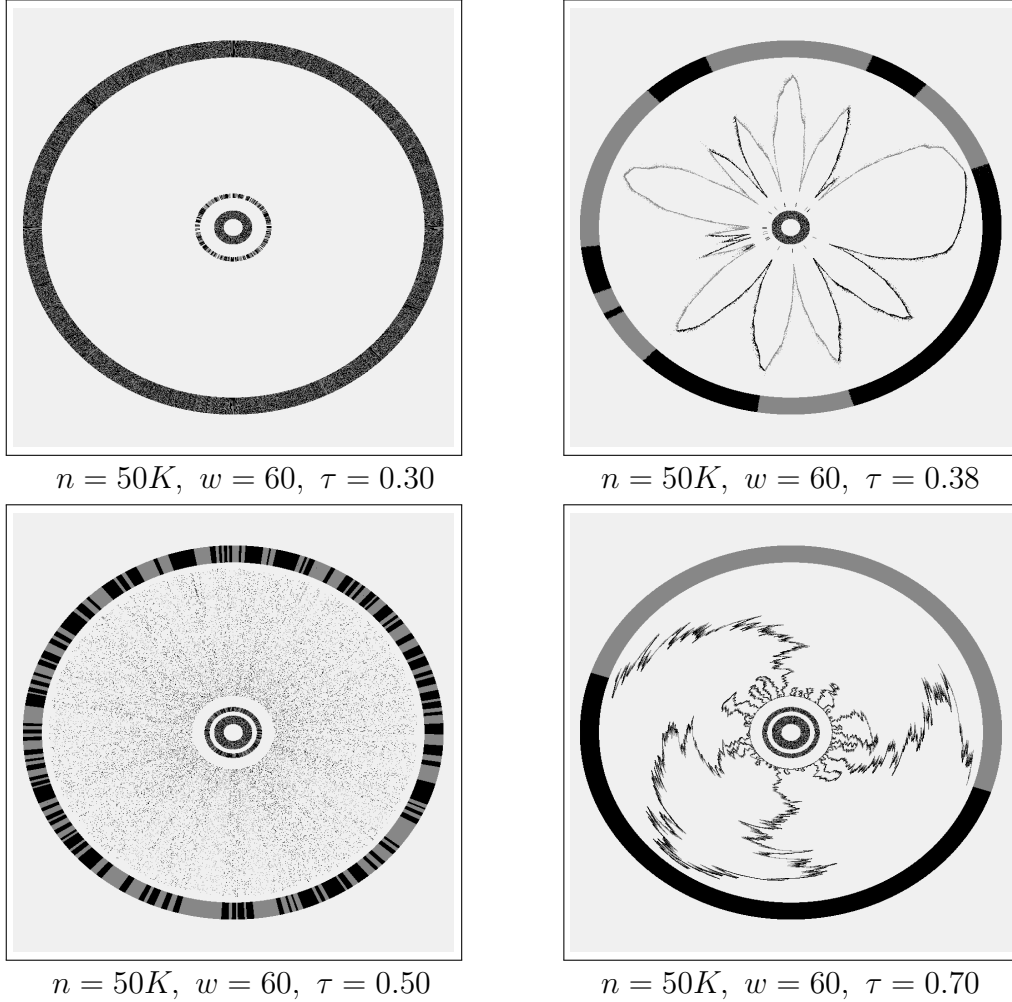


FIGURE 1. Threshold behaviour.

We shall also be interested in a variant of the model, which we shall call the *simple model*, and which proceeds in the same way except that at each stage *one* unhappy node  $u$  is selected uniformly at random and changes type so long as this does not cause it to have less nodes of its own type within  $\mathcal{N}(u)$ . The process ends when no more legal changes are possible. One might justify interest in this version of the model in two ways. Firstly there are a number of situations in which it is much easier to work with (in one instance here, and also in [2], results are proved for the standard model by first considering the simple model and then arguing that the proof can be extended to the standard case). The simple variant of the model also makes sense as soon as one drops the assumption that we are working within a closed system. One might suppose that unhappy individuals living in a city will move to a location in the same city if one should be available, but will move

elsewhere otherwise, and similarly that individuals will move into the city to fill locations becoming vacant. In all cases we shall prove the same results for both models, except for the case  $\tau > \frac{1}{2}$ , where the simple model eventually yields a society of nodes all of one type. The rough conjecture is that the simple model accurately describes local behaviour for the standard model.

In Section 2 we shall discuss terminology and we shall make some easy observations about how the evolution of the model can be understood. In Section 3 we prove Theorem 1.1. In section 4 we prove Theorem 1.2. In Section 5 we discuss the case  $\tau = \frac{1}{2}$ , which was already dealt with in [2], and in Section 6 we prove Theorem 1.4.

## 2. SOME NOTATION, TERMINOLOGY AND SOME EASY OBSERVATIONS

In describing the model earlier, we talked in terms of nodes or individuals swapping locations at various stages of the dynamic process. To work in this way, however, requires one to draw a distinction between individuals and the locations they occupy at each stage, and to maintain a bijective map at each stage between these two sets. In fact, it is notationally easier to consider a process whereby one simply has a set of  $n$  nodes, with two unhappy nodes of opposite type selected at each stage (if such exist), which may then both change type (when this occurs we shall still refer to the nodes as ‘swapping’, but now they are swapping type rather than location). Thus nodes are identified with indices for their locations amongst the set  $\{0, 1, \dots, n-1\}$ , and unless stated otherwise, addition and subtraction on these indices are performed modulo  $n$ . In the context of discussing a node  $u_1$ , for example, we might refer to the immediate neighbour on the right as node  $u_1 + 1$ . Since we work modulo  $n$  it is worth clarifying some details of the interval notation: for  $0 \leq b < a < n$ , we let  $[a, b]$  denote the set of nodes (‘interval’)  $[a, n-1] \cup [0, b]$  (while  $[b, a]$  is, of course, understood in the standard way).

As noted before, for any node  $u$ , we let  $\mathcal{N}(u)$  denote the neighbourhood of  $u$ , which is the interval  $[u-w, u+w]$ . For any set of nodes  $I$ , suppose that  $x$  is the number of  $\alpha$  nodes in  $I$ , while  $y = |I| - x$ . Then  $\Theta(I) := x - y$  and is called the *bias* of  $I$ . By the bias of a node we mean the bias of its neighbourhood. Recall that by a *run* of length  $m+1$  we mean an interval  $[u, u+m]$  in which all nodes are of the same type.

We shall be particularly interested in local configurations which are stable, in the sense that certain nodes in them can never be caused to change type. Note that if an interval of length  $w+1$  contains at least  $\tau(2w+1)$  many  $\alpha$  nodes, then each of those  $\alpha$  nodes is happy so long as the others do not change type, meaning that, in fact, no  $\alpha$  nodes in that interval will ever change type. We say that such an interval of length  $w+1$  is  $\alpha$ -stable (and similarly for  $\beta$ ). An interval of length  $w+1$  is *stable* if it is either  $\alpha$ -stable or  $\beta$ -stable. We shall also make use of a particular kind of stable interval which was used in [2]: a *firewall* is a run of length at least  $w+1$ . We write ‘for  $0 \ll w \ll n$ ’, to mean ‘for all sufficiently large  $w$  and all  $n$  sufficiently large compared to  $w$ ’.



**Arguing that the process ends.** We define the *harmony index* corresponding to any given configuration to be the sum over all nodes of the number of their own type within their neighbourhood. For  $\tau \leq \frac{1}{2}$ , this harmony index is easily seen to strictly increase whenever an unhappy node changes type, which combined with the existence of an upper bound  $n(2w + 1)$ , implies that the process must terminate after finitely many stages. For  $\tau > \frac{1}{2}$ , we shall argue that with probability tending to 1 as  $n \rightarrow \infty$ , the initial configuration is such that complete segregation eventually occurs with probability 1. Once complete segregation has occurred it is easy to see that all future states must be completely segregated, but that ‘rotations’ can occur, i.e. if the nodes of type  $\alpha$  are precisely the interval  $[a, b]$  at stage  $s$ , then at stage  $s + 1$  they must be either  $[a - 1, b - 1]$ ,  $[a, b]$  or  $[a + 1, b + 1]$ .

### 3. THE CASE $\tau < \kappa$

The analysis here is identical for the standard and simple models. Of course, we are yet to explain how  $\kappa$  comes to be defined as described previously – we shall do so shortly. The basic idea is that we wish to find the value of  $\tau$  at which stable intervals become more likely than unhappy nodes in the initial configuration. For such  $\tau$ , taking  $w$  large, we shall have that stable intervals are *much* more likely than unhappy nodes in the initial configuration. If  $u$  is selected uniformly at random then we shall almost certainly have stable intervals of both types on either side of  $u$  before any unhappy elements, meaning that  $u$  never changes type.

Recall that an interval of length  $w + 1$  is  $\beta$ -stable if it contains at least  $(2w + 1)\tau$  many nodes of type  $\beta$ . In the initial configuration this can be modelled with a binomial distribution  $X \sim b(w + 1, \frac{1}{2})$ , from which we are interested in the probability  $P_{\text{stab}} = \mathbf{P}(X \geq (2w + 1)\tau)$ . Similarly, an  $\alpha$  node is unhappy if its neighbourhood contains more than  $(2w + 1)(1 - \tau)$  many  $\beta$  nodes. We model this as  $Y \sim b(2w, \frac{1}{2})$ , from which we are interested in the probability  $P_{\text{unhap}} = \mathbf{P}(Y > (2w + 1)(1 - \tau))$ . In fact, for now we will use the approximations  $X \sim b(w, \frac{1}{2})$  where  $P_{\text{stab}} = \mathbf{P}(X \geq 2w\tau)$ , and  $Y \sim b(2w, \frac{1}{2})$  where  $P_{\text{unhap}} = \mathbf{P}(Y > 2w(1 - \tau))$  (later it will be easy to observe that these approximations suffice, and that it does not matter that we have considered the probability that an interval of  $w + 1$  is  $\beta$ -stable, rather than simply the probability that it is stable). We are interested in finding conditions on  $w$  and  $\tau$  which ensure:

$$\frac{P_{\text{unhap}}}{P_{\text{stab}}} < 1 \quad \text{or} \quad \frac{P_{\text{unhap}}}{P_{\text{stab}}} < \varepsilon.$$

Now, by the Central Limit Theorem, for large  $w$  the binomial distribution approaches the normal distribution, but, in fact, this is not enough for our purposes. We need a good estimate as to the rate at which this convergence to the normal distribution takes place. To this end, we apply a powerful result of McKay based on an estimate of Littlewood for the tail of the binomial distribution, see Theorem 2 of [15]. We use the notation  $\phi(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$ , and  $Q(x) = \int_x^\infty \phi(u)du$ . The McKay-Littlewood theorem gives :

$$P_{\text{stab}} = \frac{\sqrt{w}}{2} \cdot \binom{w-1}{h-1} \cdot \frac{1}{2^{w-1}} \cdot \frac{Q(\sqrt{w}(4\tau-1))}{\phi(\sqrt{w}(4\tau-1))} \cdot e^{E_1}$$

where  $h = \lceil 2w\tau \rceil$  and  $0 \leq E_1 \leq \text{Min} \left\{ \frac{2}{w(4\tau-1)}, \sqrt{\frac{\pi}{2w}} \right\}$ .

Similarly, we get

$$P_{\text{unhap}} = \frac{\sqrt{w}}{\sqrt{2}} \cdot \binom{2w-1}{h'-1} \cdot \frac{1}{2^{2w-1}} \cdot \frac{Q(\sqrt{2w}(1-2\tau))}{\phi(\sqrt{2w}(1-2\tau))} \cdot e^{E_2}$$

where  $h' = \lceil 2w(1-\tau) \rceil$  and  $0 \leq E_2 \leq \text{Min} \left\{ \frac{1}{w(1-2\tau)}, \frac{1}{2} \sqrt{\frac{\pi}{w}} \right\}$ . Note that for  $\tau < \frac{1}{2}$  we have  $h \leq h'$ . Now what we are really interested in is:

$$\frac{P_{\text{unhap}}}{P_{\text{stab}}} = \frac{\sqrt{2}}{2^w} \cdot \frac{\binom{2w-1}{h'-1}}{\binom{w-1}{h-1}} \cdot \frac{\phi(\sqrt{w}(4\tau-1))}{Q(\sqrt{w}(4\tau-1))} \cdot \frac{Q(\sqrt{2w}(1-2\tau))}{\phi(\sqrt{2w}(1-2\tau))} \cdot e^{E_3}$$

where  $E_3 = E_2 - E_1$  is such that  $-\sqrt{\frac{\pi}{2}} \leq E_3 \leq \frac{\sqrt{\pi}}{2}$ . We can bound each of the terms in this expression. To start with, we can use the following standard bounds: for any  $x > 0$ ,

$$\frac{2\phi(x)}{x + \sqrt{x^2 + 4}} \leq Q(x) \leq \frac{2\phi(x)}{x + \sqrt{x^2 + \frac{8}{\pi}}}.$$

For some constants  $C_1$  and  $C_2$  (which may depend on  $\tau$  but not on  $w$ ), we get:

$$\begin{aligned} C_1 \cdot \frac{1}{2^w} \cdot \frac{\binom{2w-1}{h'-1}}{\binom{w-1}{h-1}} \cdot \frac{(4\tau-1) + \sqrt{(4\tau-1)^2 + \frac{8}{w\pi}}}{\sqrt{2}(1-2\tau) + \sqrt{2(1-2\tau)^2 + \frac{4}{w}}} &\leq \frac{P_{\text{unhap}}}{P_{\text{stab}}} \\ &\leq C_2 \cdot \frac{1}{2^w} \cdot \frac{\binom{2w-1}{h'-1}}{\binom{w-1}{h-1}} \cdot \frac{(4\tau-1) + \sqrt{(4\tau-1)^2 + \frac{4}{w}}}{\sqrt{2}(1-2\tau) + \sqrt{(1-2\tau)^2 + \frac{8}{w\pi}}}. \end{aligned}$$

On the assumption that  $\frac{1}{4} < \tau < \frac{1}{2}$ , we can easily bound the fractions involving  $\tau$  and absorb these bounds into  $C_1$  and  $C_2$ , giving:

$$(3.0.1) \quad C_1 \cdot A \leq \frac{P_{\text{unhap}}}{P_{\text{stab}}} \leq C_2 \cdot A$$

where



$$A := \frac{1}{2^w} \cdot \frac{\binom{2w-1}{h'-1}}{\binom{w-1}{h-1}}.$$

To deal with  $A$ , we use the following estimates from Stirling's formula: for all positive integers  $n$ ,

$$\sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n} \leq n! \leq e n^{n+\frac{1}{2}} e^{-n}.$$

Applying these bounds we find that for constants  $D_1$  and  $D_2$ :

$$D_1 \leq A \cdot 2^w \cdot \frac{(w-1)^{w-\frac{1}{2}} (h'-1)^{h'-\frac{1}{2}} (2w-h')^{2w-h'+\frac{1}{2}}}{(2w-1)^{2w-\frac{1}{2}} (h-1)^{h-\frac{1}{2}} (w-h)^{w-h+\frac{1}{2}}} \leq D_2.$$

Now, using the fact that  $h = \lceil 2w\tau \rceil$  and  $h' = \lceil 2w(1-\tau) \rceil$  (but ignoring the ceiling issues for now) we note that the total power of  $w$  on the top and bottom of the fraction is equal to  $3w - \frac{1}{2}$ . Dividing out and absorbing constant terms into  $D_1$  and  $D_2$  we get:

$$D_1 \leq A \cdot 2^w \cdot \frac{(1 - \frac{1}{w})^{w-\frac{1}{2}} (2(1-\tau) - \frac{1}{w})^{2w(1-\tau)-\frac{1}{2}} (2\tau)^{2w\tau+\frac{1}{2}}}{(2 - \frac{1}{w})^{2w-\frac{1}{2}} (2\tau - \frac{1}{w})^{2w\tau-\frac{1}{2}} (1-2\tau)^{w(1-2\tau)+\frac{1}{2}}} \leq D_2,$$

and

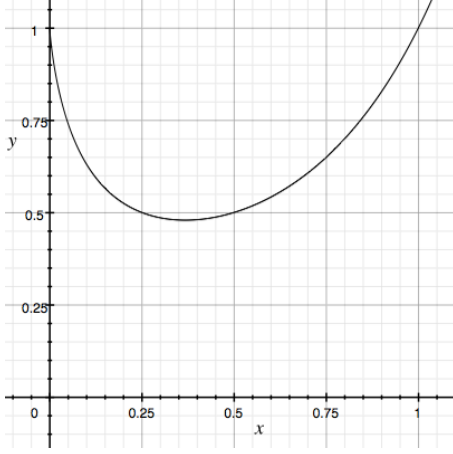
$$D_1 \leq A \cdot 2^w \cdot \frac{(1 - \frac{1}{w})^{w-\frac{1}{2}} 2^{2w-2w\tau-\frac{1}{2}} (1-\tau)^{2w-2w\tau-\frac{1}{2}} (1 - \frac{1}{2w(1-\tau)})^{2w(1-\tau)-\frac{1}{2}} (2\tau)^{2w\tau+\frac{1}{2}}}{2^{2w-\frac{1}{2}} (1 - \frac{1}{2w})^{2w-\frac{1}{2}} (2\tau)^{2w\tau-\frac{1}{2}} (1 - \frac{1}{2w\tau})^{2w\tau-\frac{1}{2}} (1-2\tau)^{w(1-2\tau)+\frac{1}{2}}} \leq D_2.$$

For  $x \geq 2$ ,  $e^{-1} \leq (1 - \frac{1}{x})^{x-\frac{1}{2}} \leq 1$ . Furthermore, the reintroduction of the ceiling function introduces no difficulty here, as for  $0 \leq \delta < 1$  representing  $\lceil x \rceil - x$  we also have  $e^{-1} \leq (1 + \frac{\delta-1}{x})^{x-\frac{1}{2}} \leq 1$ . Thus:

$$(3.0.2) \quad D_1 \leq A \cdot \left( \frac{2^{1-2\tau} \cdot (1-\tau)^{2(1-\tau)}}{(1-2\tau)^{1-2\tau}} \right)^w \leq D_2.$$

Combining 3.0.1 and 3.0.2 and merging constant terms, we get:

$$(3.0.3) \quad C_1 \cdot \left( \frac{(\frac{1}{2} - \tau)^{1-2\tau}}{(1-\tau)^{2(1-\tau)}} \right)^w \leq \frac{P_{\text{unhap}}}{P_{\text{stab}}} \leq C_2 \cdot \left( \frac{(\frac{1}{2} - \tau)^{1-2\tau}}{(1-\tau)^{2(1-\tau)}} \right)^w.$$

FIGURE 2.  $y = x^{2x}$ 

Now put  $f(x) = x^{2x}$ . Then the value  $\kappa$  we are looking for is exactly  $\tau$  such that

$$f\left(\frac{1}{2} - \tau\right) = f(1 - \tau).$$

In other words  $x = \frac{1}{2} - \kappa$  is exactly such that

$$f(x) = f\left(x + \frac{1}{2}\right).$$

Since  $f$  has a unique turning point at  $e^{-1}$ , it follows that  $\kappa$  is unique, and numerical analysis gives

$$\kappa \approx 0.353092313$$

(which is just slightly less than  $\sqrt{2}/4$ ).

Now from (3.0.3) it is clear that these conclusions are unaffected by our use of the approximations  $X \sim b(w, \frac{1}{2})$  where  $P_{\text{stab}} = \mathbf{P}(X \geq 2w\tau)$ , and  $Y \sim b(2w, \frac{1}{2})$  where  $P_{\text{unhap}} = \mathbf{P}(Y > 2w(1 - \tau))$ . Since, for fixed  $\tau$ , the *actual* probabilities that in the initial configuration a given interval of length  $w + 1$  is  $\beta$ -stable or that a given node is unhappy, are to within a fixed multiplicative constant of these approximations, (3.0.3) remains true when  $P_{\text{unhap}}$  and  $P_{\text{stab}}$  are replaced by these actual probabilities. The analysis above also considers a specific interval of length  $w + 1$  to which a node might belong and which might be stable, and so does not overtly take account of the fact that a given node belongs to  $w + 1$  many intervals of length  $w + 1$  which might be stable. Again, (3.0.3) shows that this does not matter – when  $\tau < \kappa$  and  $k, r > 0$ , for all sufficiently large  $w$  if  $u$  is chosen uniformly at random then, in the initial configuration, it is more than  $kw^r$  times as likely that  $u$  belongs to a stable interval than that it is an unhappy node. When  $\tau > \kappa$  the reverse is true. We need something more than this though. For any  $u$ , let  $x_u$  be the first node to the left of  $u$  which, in the initial configuration, is either unhappy or else belongs to an interval which is stable. We must show that, for any  $\epsilon' > 0$ , if  $0 \ll w \ll n$  (i.e. if  $w$  is large and  $n$  is large compared to  $w$ ) then the following occurs with probability  $> 1 - \epsilon'$  for  $u$  chosen uniformly at random:  $x_u$  is defined and no node in  $[x_u - 2w, x_u]$  is unhappy in the initial configuration. From this we can easily argue that  $u$  selected uniformly at random will almost certainly have stable intervals of both types on either side in the initial configuration before any unhappy nodes, as required. The following lemma (which is stated in such a way as to be general purpose, so we can apply it again later) allows us to do this:

**Lemma 3.1.** *Let  $P_u$  and  $Q_u$  be events which only depend on the neighbourhood of  $u$  in the initial configuration, meaning that if the neighbourhood of  $v$  in the initial configuration is*

identical that of  $u$  (i.e. for all  $i \in [-w, w]$ ,  $u + i$  is of the same type as  $v + i$ ), then  $P_u$  holds iff  $P_v$  holds and  $Q_u$  holds iff  $Q_v$  holds. Suppose also that:

- (i)  $\mathbf{P}(P_u) \neq 0$  and  $\mathbf{P}(Q_u) \neq 0$ .
- (ii) For all  $k$ , for all sufficiently large  $w$ ,  $\mathbf{P}(P_u)/\mathbf{P}(Q_u) > kw$ .

For any  $u$ , let  $x_u$  be the first node to the left<sup>1</sup> of  $u$  such that either  $P_u$  or  $Q_u$  holds. For any  $\epsilon > 0$ , if  $0 \ll w \ll n$  then the following occurs with probability  $> 1 - \epsilon$  for  $u$  chosen uniformly at random:  $x_u$  is defined and for no node  $v$  in  $[x_u - 2w, x_u]$  does  $Q_v$  hold.

An analogous result holds when ‘left’ is replaced by ‘right’.

**Proof.** For a general node  $u$ , we’ll say it is of *type 1* if  $P_u$  holds and no node  $u' \in [u - 2w, u]$  satisfies  $Q_{u'}$ , and of *type 2* if  $Q_u$  holds. Let  $\pi$  be the probability, for  $u$  chosen uniformly at random, that  $x_u$  is defined and of type 1. Our aim is to show that  $\pi > 1 - \epsilon$ , for all  $0 \ll w \ll n$ . We define an iteration which assigns colours to nodes as follows.

Step 0. Pick a node  $t_0$  uniformly at random.

Step  $s + 1$ . Let  $v_s$  be the first node to the left of  $t_s$ , such that  $v_s = x_{t_s}$  or such that  $s > 0$  and  $v_s = t_0$ . Carry out the instructions for the first case below which applies:

- (1) If there exists no such  $v_s$  then terminate the iteration.
- (2) If  $v_s = t_0$  and  $s > 0$  then make  $t_s$  undefined and terminate the iteration.
- (3) If there exist any nodes  $x$  in  $[v_s - 2w, v_s]$  such that  $Q_x$  holds, then colour  $t_s$  black, otherwise it is of type 1 and we colour it white. Define  $t_{s+1} = v_s - (2w + 1)$ , unless  $t_0$  lies in the interval  $[v_s - (2w + 1), v_s]$ , in which case terminate the iteration.

This completes the description of the iteration. Let  $S$  be the maximum value of  $s$  for which  $t_s$  is defined and coloured.

First note that hypothesis (i) in the statement of the lemma guarantees that  $S \rightarrow \infty$  as  $n \rightarrow \infty$ . Similarly, we may assume that at least one  $t_s$  is coloured black. Now let  $\pi'$  be the proportion of the  $t_s$  which are coloured white, and observe that with high probability  $\pi' \rightarrow \pi$  as  $n \rightarrow \infty$ , i.e. for all  $\epsilon' > 0$ , for all  $0 \ll n$ , we have  $|\pi - \pi'| < \epsilon'$  with probability  $> 1 - \epsilon'$ . In order to see this, consider the situation at the beginning of step  $s + 1$  of the iteration, when  $s > 0$ . In order to define  $t_s$ , we moved left until finding  $x_{t_{s-1}}$ , and then moved a further  $2w + 1$  nodes to the left. So far, then, nothing that has happened in the iteration tells us anything about the neighbourhood of  $t_s$  and all those nodes to the left of  $t_s$  and strictly to the right of  $t_0$ . So long as  $S > s$ , the way in which  $t_s$  is coloured depends only on these nodes.

So, we now wish to show that for all sufficiently large  $w$ , with probability tending to 1 as  $n \rightarrow \infty$ , we have  $\pi' > 1 - \epsilon$ . Let  $\rho$  be the ratio of type 1 nodes to type 2 nodes (in  $[0, n - 1]$ ), and consider the interval  $(t_{s+1}, t_s]$ . However  $t_s$  is coloured, we have at most  $(2w + 1)$  many type 1 nodes in here, namely some subset of  $[v_s - 2w, v_s]$ . Similarly, if  $t_s$

<sup>1</sup>By the first node to the left of  $u$  satisfying a certain condition we mean the first in the sequence  $u, u - 1, u - 2, \dots$  which satisfies the condition.

is coloured black then we get at least one type 2 node in here. Thus, summing over all intervals  $(t_{s+1}, t_s]$ , we find that

$$\rho \leq \frac{(2w+1)S}{|\{s \leq S : t_s \text{ is black}\}|}$$

which is to say

$$\rho \leq \frac{2w+1}{1-\pi'}.$$

So if we can show that  $\rho \gg 2w+1$ , it will follow that  $\pi'$  is close to 1 as required. Let  $p_1$  and  $p_2$  be the probabilities that uniformly randomly selected  $u$  itself is of type 1 and 2 respectively. Then, by the law of large numbers, we have  $\left|\rho - \frac{p_1}{p_2}\right| < \epsilon'$  with probability approaching 1 as  $n \rightarrow \infty$ . Now property (ii) from the statement of the lemma means that for each  $k$ , for all sufficiently large  $w$ ,  $\frac{p_1}{p_2} > kw$ , giving the result.  $\square$

We now want to use Lemma 3.1 to show:

- (†<sub>1</sub>) For any  $\epsilon > 0$ , for  $0 \ll w \ll n$ , if  $u$  is selected uniformly at random, then with probability  $> 1 - \epsilon$ , there exist  $u_2 < u_1 < u$  such that  $u_1$  and  $u_2$  belong to stable intervals of opposite types,  $|u_1 - u_2| > 2w$ ,  $|u - u_1| > 2w$  and there are no unhappy nodes in the interval  $[u_2, u]$ .

(Of course a similar result will then hold to the right of  $u$ .) Choose  $\epsilon' \ll \epsilon$ . For any  $v$ , let  $P_v$  be the event that  $v$  belongs to a stable interval in the initial configuration, let  $Q_v$  be the event that  $v$  is an unhappy node in the initial configuration and let  $x_v$  be the first node  $u$  to the left of  $v$  such that either  $P_u$  or  $Q_u$  holds. By Lemma 3.1 and (3.0.3), since  $\tau < \kappa$  we may let  $w$  be large enough (and  $n$  large enough compared to  $w$ , as always) such that the following occurs with probability  $> 1 - \epsilon'$  for  $u$  chosen uniformly at random:  $x_u$  is defined and for no node  $v$  in  $[x_u - 2w, x_u]$  does  $Q_v$  hold. Also, let  $\epsilon'' < \epsilon'$  be such that  $1 - \epsilon''$  is exactly the probability, for  $u$  chosen uniformly at random, that  $x_u$  is defined and for no node  $v$  in  $[x_u - 2w, x_u]$  does  $Q_v$  hold. We inductively define a sequence of nodes, much as in the proof of Lemma 3.1.

Step 0. Pick a node  $u$  uniformly at random. If  $x_u$  is undefined, or  $u \in [x_u - 2w, x_u]$ , or if any node  $z$  in  $[x_u - 2w, x_u]$  satisfies  $Q_z$ , then declare that the iteration has terminated unsuccessfully. Otherwise, let  $y = x_u - (2w + 1)$  and consider  $x_y$ . If  $x_y$  is undefined, or  $u \in [x_y - 2w, x_y]$ , or if any node  $z$  in  $[x_y - 2w, x_y]$  satisfies  $Q_z$ , then declare that the iteration has terminated unsuccessfully. Otherwise let  $\gamma$  be the type of a stable interval containing  $x_y$ , define  $u_1 = x_y$  and  $t_0 = x_y - 2w + 1$ .

Step  $s + 1$ . Let  $v$  be the first node to the left of  $t_s$ , such that  $v = x_{t_s}$  or such that  $v = u$ . Carry out the instructions for the first case below which applies:

- (1) If  $u \in [v - 2w, v]$  then declare that the iteration has terminated unsuccessfully.
- (2) If there exist any nodes  $z$  in  $[v - 2w, v]$  such that  $Q_z$  holds, then declare that the iteration has terminated unsuccessfully. Otherwise, if  $v$  belongs to a stable interval which is not of type  $\gamma$  then define  $u_2 = v$  and declare that the iteration

has terminated successfully. If the previous cases do not apply then define  $t_{s+1} = v - (2w + 1)$  and proceed to the next step of the iteration.

Note first that if the iteration terminates successfully then  $u_1$  and  $u_2$  exist as required. Let us say that the iteration ends prematurely if either of  $x_u$  or  $x_y$  are undefined at step 0, or if  $u \in [x_u - 2w, x_u] \cup [x_y - 2w, x_y]$ , or else there exists  $i > 0$  such that the iteration has not terminated strictly prior to step  $i$  and case (1) applies at step  $i$  (so, roughly, the iteration ends prematurely if it runs out of nodes before finding either an unhappy node or both  $u_1$  and  $u_2$ ). Letting  $\zeta$  be the probability that the iteration ends prematurely, note that  $\zeta \rightarrow 0$  as  $n \rightarrow \infty$  (for fixed  $w$ ). Then (for  $\epsilon''$  as specified above) the probability that the iteration terminates unsuccessfully is less than:

$$\zeta + \epsilon' + (1 - \epsilon'')\epsilon' + (1 - \epsilon'')^2\epsilon' + (1 - \epsilon'')^3\frac{1}{2}\epsilon' + (1 - \epsilon'')^4\frac{1}{2^2}\epsilon' + \dots$$

So for sufficiently large  $n$ , the probability the iteration terminates unsuccessfully is less than  $4\epsilon'$ , establishing  $(\dagger_1)$ .

**Definition 3.2.** For  $\gamma \in \{\alpha, \beta\}$ , let  $\gamma^*$  be the opposite type, i.e.  $\gamma^* = \alpha$  if  $\gamma = \beta$ , and  $\gamma^* = \beta$  otherwise.

To complete the argument, we now show that the existence of stable intervals of opposite types on both sides before any unhappy nodes, suffices to prevent type changes, i.e. that  $(\dagger_1)$  and its mirror image give the desired result. So suppose that  $u_0$  satisfies the property that, in the initial configuration, there exist  $u_{-2} < u_{-1} < u_0 < u_1 < u_2$  such that:

- For  $i \in [-2, 1]$ ,  $|u_{i+1} - u_i| > 2w$ .
- For  $i \in [-2, 2] - \{0\}$ ,  $u_i$  belongs to a  $\gamma_i$ -stable interval,  $I_i$  say. Also  $\gamma_{-2} = \gamma_{-1}^*$  and  $\gamma_2 = \gamma_1^*$ .
- There are no unhappy nodes in the interval  $[u_{-2}, u_2]$ .

For  $i \in [-2, 2] - \{0\}$ , let  $I_i = [a_i, b_i]$ . Then it follows directly by induction on stages, that nodes in the interval  $[a_{-2}, a_{-1})$  can only change from  $\gamma_{-1}$  to  $\gamma_{-2}$ , that nodes in the interval  $[a_{-1}, b_1]$  cannot change type, and that nodes in the interval  $(b_1, b_2]$  can only change from  $\gamma_1$  to  $\gamma_2$ .

#### 4. THE CASE $\kappa < \tau < \frac{1}{2}$

We work first with the simple model, and then supply the necessary modifications for the standard case. In what follows we shall work with some fixed  $\tau$  in the interval  $(\kappa, \frac{1}{2})$ , some fixed  $\epsilon > 0$ , and we shall assume that  $n$  is large compared to  $w$ . We want to show that there exists a constant  $d$  such that for all sufficiently large  $w$  the probability that a randomly chosen element will belong to a run of length  $\geq e^{w/d}$  is greater than  $1 - \epsilon$ . Of course, proving the result for all sufficiently large  $w$  suffices to give the result for all  $w$  since one can simply adjust the choice of  $d$  to deal with finitely many small values, but we shall make frequent use of the fact that we need only work for all sufficiently large  $w$ .

in what follows and so stating the theorem in this way is instructive.

Our entire analysis takes place relative to a node  $u_0$  which is chosen uniformly at random. In order to specify how the type changing process can be expected to develop in the vicinity<sup>2</sup> of  $u_0$ , we shall detail a finite number of events  $Q_i$ , such as a lack of unhappy nodes in the neighbourhood of  $u_0$  in the initial configuration, for example. Our aim is to show that, for any  $\epsilon' > 0$ , the probability that all the  $Q_i$  occur is greater than  $1 - \epsilon'$  for all sufficiently large  $w$  (and for  $n$  sufficiently large compared to  $w$ ). Suppose that this is established for the finite set of events  $Q_i$  with  $i \in \Pi$  (for some finite set  $\Pi$ ). To establish the result when a new event  $Q_j$  with  $j \notin \Pi$  is included, it then suffices to prove for each  $\epsilon' > 0$  and all sufficiently large  $w$ , that the probability of  $Q_j$  given  $P$  is greater than  $1 - \epsilon'$ , where  $P$  is any conjunction (possibly empty) of the  $Q_i$  such that  $i \in \Pi$ . Of course, we choose that  $P$  which is most convenient to work with.

Recall that for any set of nodes  $I$ ,  $\Theta(I)$  is the *bias* of  $I$  and that by the bias of a node we mean the bias of its neighbourhood.

**The initial configuration.** Let us begin by considering what can be expected from the initial configuration in the vicinity of  $u_0$ . In general, if  $x_1, \dots, x_k$  are independent random variables with  $\mathbf{P}(x_i = 1) = \mathbf{P}(x_i = -1) = \frac{1}{2}$  when  $1 \leq i \leq k$ , then letting  $X = \sum_{i=1}^k x_i$  Hoeffding's inequality gives, for arbitrary  $\lambda > 0$ :

$$\mathbf{P}(|X| > \lambda\sqrt{k}) < 2e^{-\lambda^2/2}.$$

Now we use this to bound the probability that a node  $u$  has bias in the initial configuration which will cause it to be unhappy, should  $u$  be of the minority type in its neighbourhood. So, we wish to bound the probability that the number of  $\alpha$  nodes in  $\mathcal{N}(u)$  is  $> (1 - \tau)(2w + 1)$  or the number of  $\beta$  nodes in  $\mathcal{N}(u)$  is  $> (1 - \tau)(2w + 1)$ . This corresponds to a bias  $\Theta(\mathcal{N}(u))$  of  $> (1 - 2\tau)(2w + 1)$  or  $< -(1 - 2\tau)(2w + 1)$ .

**Definition 4.1.** When  $\Theta(\mathcal{N}(u)) > (1 - 2\tau)(2w + 1)$  or  $\Theta(\mathcal{N}(u)) < -(1 - 2\tau)(2w + 1)$  we say that  $u$  has **high bias**, denoted  $\mathbf{Hb}(u)$ . If this holds for the initial configuration, we say that  $\mathbf{Hb}^*(u)$  holds.

**Definition 4.2.** For the remainder of this section we define  $d = \frac{2}{(1-2\tau)^2}$ .

**Lemma 4.3** (Likely happiness). *Let  $u$  be a node chosen uniformly at random. For any  $\epsilon' > 0$  and for all sufficiently large  $w$ , the probability that  $\mathbf{Hb}^*(u)$  holds is  $< \epsilon'e^{-w/d}$ .*

**Proof.** Putting  $\lambda\sqrt{2w + 1} = (1 - 2\tau)(2w + 1)$  in Hoeffding's inequality above, we get

$$\lambda^2/2 = (1 - 2\tau)^2 \frac{(2w + 1)}{2},$$

---

<sup>2</sup>To be clear, the term ‘vicinity’ of  $u_0$  is used informally here, to mean, roughly, an interval containing  $u_0$ , which may be large compared to  $w$  but which is small compared to  $n$ .

so

$$e^{-\lambda^2/2} = e^{-w/d^*} \quad \text{where} \quad d^* = \frac{2w}{2w+1} \cdot \frac{1}{(1-2\tau)^2}.$$

We chose  $d > \frac{1}{(1-2\tau)^2}$ , which means that for any  $\epsilon' > 0$  and for all sufficiently large  $w$ , the probability that  $u$  has high bias in the initial configuration is  $< \epsilon' e^{-w/d}$ , as required.  $\square$

**Defining the nodes  $l_i$  and  $r_i$ .** For now, we fix some  $k_0 > 0$  (we shall choose a specific value of  $k_0$  which is appropriate later) and for  $1 \leq i \leq k_0$  we define a node  $l_i$  to the left of  $u_0$  and also a node  $r_i$  to the right. We let  $l_1$  be the first node  $v$  to the left of  $u_0$  such that  $\text{Hb}^*(v)$  holds, so long as this node is in the interval  $[u_0 - \frac{1}{4}n, u_0]$  (otherwise  $l_1$  is undefined). Then, given  $l_i$  for  $i < k_0$  we let  $l_{i+1}$  be the first node  $v$  to the left of  $l_i - (2w+1)$  such that  $\text{Hb}^*(v)$  holds, so long as no nodes in the interval  $[l_{i+1}, l_i]$  are outside the interval  $[u_0 - \frac{1}{4}n, u_0]$  (otherwise  $l_{i+1}$  is undefined). We let  $r_1$  be the first node  $v$  to the right of  $u_0$  such that  $\text{Hb}^*(v)$  holds, so long as this node is in the interval  $[u_0, u_0 + \frac{1}{4}n]$ . Given  $r_i$  for  $i < k_0$  we let  $r_{i+1}$  be the first node  $v$  to the right of  $r_i + (2w+1)$  such that  $\text{Hb}^*(v)$  holds, so long as no nodes in the interval  $[r_i, r_{i+1}]$  are outside the interval  $[u_0, u_0 + \frac{1}{4}n]$ .

The reason for considering the intervals  $[u_0 - \frac{1}{4}n, u_0]$  and  $[u_0, u_0 + \frac{1}{4}n]$  in the above, is that we wish to be able to move left from  $u_0$  to  $l_{k_0}$  without meeting any of the nodes  $r_i$ .

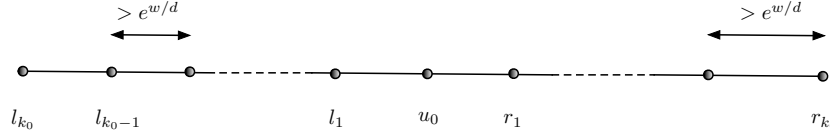


FIGURE 3. Picking out nodes of high bias in the vicinity of  $u_0$ .

**Definition 4.4** (Defining  $Q_0$ ). *Let  $d$  be as in Definition 4.2. It is notationally convenient to let  $l_0 = u_0 = r_0$ . We let  $Q_0$  be the event that for  $1 \leq i \leq k_0$ :*

- (i)  $l_i$  and  $r_i$  are both defined, and;
- (ii)  $|l_i - l_{i-1}| > e^{w/d}$  and  $|r_i - r_{i-1}| > e^{w/d}$ .

Note that for any fixed  $w$ , as  $n \rightarrow \infty$  the probability that any  $l_i$  or  $r_i$  is undefined (for  $i \leq k_0$ ) goes to 0. By Lemma 4.3, and since the probability that any node in an interval  $I$  has high bias in the initial configuration is at most  $\sum_{u \in I} \mathbf{P}(\text{Hb}^*(u))$ , for any  $\epsilon' > 0$  and for any fixed  $k_0 \geq 1$  we can ensure that  $\mathbf{P}(Q_0) > 1 - \epsilon'$  by taking  $w$  sufficiently large (and by taking  $n$  sufficiently large compared to  $w$ ). Thus, for  $0 \ll w \ll n$ , the picture we are presented with is almost certainly as in Figure 3.

So far we have defined an event,  $Q_0$ , which ensures that for  $1 \leq i \leq k_0$ , each  $l_i$  and each  $r_i$  is defined and that these nodes are at very large distances from each other for large  $w$ .



The way in which we defined each  $l_i$  and  $r_i$  means that no node in the interval  $[l_{k_0}, r_{k_0}]$  can be unhappy in the initial configuration unless it is in one of the intervals  $[l_i - 2w, l_i]$  or  $[r_i, r_i + 2w]$ . We are yet to choose  $k_0$ .

**Building the informal picture.** Recall that, by a *run* of length  $m + 1$  we mean an interval  $[u, u + m]$  in which all nodes are of the same type, and that a *firewall* is a run of length at least  $w + 1$ .

Let us now consider informally what can be expected to happen in the neighbourhood of  $l_i$ . Suppose that  $l_i$  is initially of type  $\beta$  and is unhappy in the initial configuration. Then with high probability for sufficiently large  $w$ , there will not be any unhappy nodes of type  $\alpha$  in the neighbourhood of  $l_i$  in the initial configuration. If  $l_i$  changes type, then this will make the bias in its neighbourhood still more positive, which may cause further nodes of type  $\beta$  to become unhappy. If these change to type  $\alpha$  then this will further increase the bias, potentially causing more nodes to become unhappy, and so on. The following definitions formalise some of the ways in which this process might play out, and in particular the possibility that this process might play out without interference from what happens in other neighbourhoods  $\mathcal{N}(l_j)$  or  $\mathcal{N}(r_j)$ .

**Definition 4.5.** For  $0 < i < k_0$  we say that  $l_i$  **completes** at stage  $s$  if both:

- (1) No node in  $\mathcal{N}(l_i)$  is unhappy at stage  $s$ , and this is not true for any  $s' < s$ .
- (2) There exist  $x_0$  and  $x_1$  with  $l_{i+1} + 2w < x_0 < l_i - 2w < l_i + 2w < x_1 < l_{i-1} - 2w$ , such that by the end of stage  $s$ , no node in  $[x_0 - w, x_0]$  or  $[x_1, x_1 + w]$  has changed type.

We say that  $l_i$  **completes** if it completes at some stage. We also define completion for  $r_i$  analogously.

**Definition 4.6.** We say that  $l_i$  (or  $r_i$ ) **originates a firewall** if it completes at some stage  $s$  and belongs to firewall at that stage.<sup>3</sup>

The informal idea, is that we now wish to show that each  $l_i$  and each  $r_i$  has some reasonable chance of originating a firewall. Then we can choose  $k_0$  so that the probability none of the  $l_i$  originate a firewall or none of the  $r_i$  originate a firewall is  $\ll \epsilon$ , i.e. with high probability firewalls will originate either side of  $u_0$  within the interval  $[l_{k_0}, r_{k_0}]$ . Then, letting  $i_1$  be the least  $i$  such that  $l_i$  originates a firewall, and letting  $i_2$  be the least  $i$  such that  $r_i$  originates a firewall, we wish to show that with high probability the firewalls originated at  $l_{i_1}$  and  $r_{i_2}$  will spread until  $u_0$  is contained in one of them. Since these two firewalls have originated at nodes which are at distance at least  $e^{w/d}$  apart,  $u_0$  ultimately belongs to a firewall of at least this length.

<sup>3</sup>This definition might initially seem to neglect the possibility, for example, that  $l_i$  completes at some stage and does not belong to a firewall at that stage, but that, nevertheless, the sequence of type changes in its neighbourhood and surrounding neighbourhoods which have led to completion have caused the creation of a firewall. In fact we shall be able to ensure (Lemma 4.25) that with high probability such events do not occur.

In order to make this basic picture work, however, we need to be clear about the way in which firewalls will spread. If the interval  $[u - w, u]$  is a firewall of type  $\alpha$ , then when  $u + 1$  is of type  $\beta$ , it cannot be happy unless the interval  $[u + 1, u + w + 1]$  is  $\beta$ -stable. So firewalls will spread until they hit stable intervals of the opposite type. Now suppose that, with  $i_1$  and  $i_2$  as above,  $i_1 = i_2 = 2$  and, for now, suppose that  $\alpha$ -firewalls are originated at both  $l_2$  and  $r_2$ . In order to show that these two firewalls will spread until they meet each other, it will be helpful first of all, to be able to assume that in the initial configuration there are no stable subintervals of  $[l_{k_0}, r_{k_0}]$ . This will follow quite easily for large  $w$ , from our previous analysis of the ratio between the probability of unhappy nodes and stable intervals. A further danger that we have to be able to avoid, however, is that, while  $l_1$  and  $r_1$  do not originate firewalls, they *do* get as far as creating  $\beta$ -stable intervals.

**Definition 4.7.** *Given  $i$  with  $0 < i < k_0$ , let  $u = l_i$  or  $u = r_i$ . Let  $u_1 = u - (2w + 1)$  and  $u_2 = u + (2w + 1)$ . We say that  $u$  **subsides** if it completes at some stage  $s$ , and:*

- *There are no nodes in  $[u_1, u_2]$  belonging to stable intervals at stage  $s$ ;*
- *No nodes in  $\mathcal{N}(u_1)$  or  $\mathcal{N}(u_2)$  have changed type by the end of stage  $s$ .*

So we need to be able to show, in fact, that with high probability each  $l_i$  and  $r_i$  either originates a firewall or subsides. To do this clearly involves analysing what is likely to happen in each of the neighbourhoods  $\mathcal{N}(l_i)$  and  $\mathcal{N}(r_i)$ . First of all, the large distances between these nodes mean that, for fixed  $k_0$  and sufficiently large  $w$ , we can expect all of the  $l_i$  and  $r_i$  (for  $0 < i < k_0$ ) to complete, so that one can understand the early stages of the process for each of these neighbourhoods by considering each in isolation. For each of these neighbourhoods we then wish to show that a certain dichotomy holds: either a small number of type changes will occur before completion, or else a large number of type changes will occur before completion and a firewall will be created. Now, if we strengthen our original requirement that there are no stable subintervals of  $[l_{k_0}, r_{k_0}]$  in the initial configuration, to a requirement that there are no subintervals which are ‘close’ to being stable in the initial configuration (where ‘close’ is to be made precise in such a way as to ensure that when a small number of type changes occur in the neighbourhood of  $l_i$  before completion, these are not enough to create any stable intervals), then we shall have that with high probability each  $l_i$  and each  $r_i$  either originates a firewall or else completes without creating stable intervals.

Once all this is in place, there is then one further hurdle. In the above, we assumed that the firewalls originating at  $l_2$  and  $r_2$  are both  $\alpha$ -firewalls. If they are firewalls of opposite type, however, we still have some work to do in order to prove that  $u_0$  will almost certainly end up belonging to one of these two firewalls.

**Formalising the intuitive picture.** We now wish to define further events which ensure that the neighbourhoods of each  $l_i$  and  $r_i$  are nicely behaved in the initial configuration, in order that we can establish the dichotomy described above, regarding the number of changes that can be expected in the neighbourhood of each  $l_i$  and  $r_i$  prior to completion.

The distributions for these intervals in the initial configuration, however, are a little difficult to attack directly because of the nature of their definition. In choosing  $l_1$  we move left until we find the *first* node which has high bias – this gives an asymmetry to the given information concerning the neighbourhood. Roughly, we might expect something like a hypergeometric distribution, but how good is this as an approximation? What we shall do, in fact, is first of all to understand what can be expected from the neighbourhood of a node which is chosen uniformly at random from among those with borderline bias:

**Definition 4.8.** *Let us say that a node  $u$  has **borderline bias**, denoted  $\text{Bb}(u)$ , if:*

$$|\Theta(\mathcal{N}(u))| = \text{Min}\{\theta \in 2\mathbb{N} + 1 : \theta > (1 - 2\tau)(2w + 1)\},$$

*i.e.,  $u$  has high bias but decreasing the modulus of the bias by the minimum possible amount of 2 would cause it not to have high bias. We say that  $\text{Bb}^*(u)$  holds if  $u$  has borderline bias in the initial configuration.*

Note that each of the nodes  $l_i$  and  $r_i$  for  $0 < i \leq k_0$  has borderline bias.

In what follows it is often convenient to work with some fixed  $k \geq 1$  and to divide an interval  $I = [a, b]$  into  $k$  parts of equal length. This occasions the minor inconvenience that the length of the interval might not be a multiple of  $k$ , motivating the following definition:

**Definition 4.9.** *Let  $I = [a, b]$  and suppose  $k \geq 1$ . We define the subintervals  $I(1 : k) := [a, a + \lfloor \frac{b-a}{k} \rfloor] := [I(1 : k)_1, I(1 : k)_2]$  and*

$$I(j : k) := \left[ a + \left\lfloor \frac{(j-1)(b-a)}{k} \right\rfloor + 1, a + \left\lfloor \frac{j(b-a)}{k} \right\rfloor \right] := [I(j : k)_1, I(j : k)_2]$$

*for  $2 \leq j \leq k$ .*

In Definition 4.9 the intervals are counted from left to right, but it is also useful sometimes to work from right to left:

**Definition 4.10.** *Let  $I = [a, b]$  and suppose  $k \geq 1$ . For  $1 \leq j \leq k$  we define  $I(j : k)^- = I(k - j + 1 : k)$ ,  $I(j : k)_1^- = I(k - j + 1 : k)_1$  and  $I(j : k)_2^- = I(k - j + 1 : k)_2$ .*

**Lemma 4.11** (Smoothness Lemma). *Suppose  $u$  is such that the proportion of  $\alpha$  nodes in  $I := \mathcal{N}(u)$  is  $\theta$ , and that  $u$  is selected uniformly at random from nodes with this property. Then for any fixed  $k \geq 1$  and  $\epsilon' > 0$ , for all sufficiently large  $w$  the following holds with probability  $> 1 - \epsilon'$ : for every  $j$  with  $1 \leq j \leq k$ , the proportion of the nodes in  $I(j : k)$  which are  $\alpha$  nodes, lies in the interval  $[\theta - \epsilon', \theta + \epsilon']$ .*

**Proof.** Once we are given that  $\theta$  is the proportion of the nodes in  $I$  which are of type  $\alpha$ , the nodes in this interval cease to be independently distributed, and the distribution becomes hypergeometric. Since we consider fixed  $k$  and  $\epsilon'$  and take  $w$  large, it suffices to prove the result for given  $j$  with  $1 \leq j \leq k$ , i.e. if  $P_1, \dots, P_k$  are events, each of which occurs with probability tending to 1 as  $w \rightarrow \infty$ , then their conjunction also occurs with

probability tending to 1 as  $w \rightarrow \infty$ . For a given  $j$ , the result follows directly, however, from an application of Chebyshev's inequality and standard results for the mean and variance of a hypergeometric distribution. Let  $x$  be the number of  $\alpha$  nodes in the interval  $I(j : k)$  and let  $\ell$  be the length of the interval, so that  $|\ell - (2w + 1)/k| \leq 1$ . Then we have:

$$\mathbf{P}(|x/\ell - \theta| > \epsilon') < \ell^{-2} \epsilon'^{-2} \text{Var}(x) = O(1) \ell^{-1}.$$

□

Lemma 4.11 basically tells us that if we choose a node  $u$  with borderline bias uniformly at random, then for large  $w$  we can expect the bias to move towards 0 fairly smoothly as we move to  $u + (2w + 1)$  or  $u - (2w + 1)$ . In order to see roughly why this is true, suppose that  $|\Theta(\mathcal{N}(u))| = \rho$  and let  $\theta$  be the proportion of the nodes in  $\mathcal{N}(u)$  which are of type  $\alpha$ . Let  $I = [u, u + (2w + 1)]$  and, for some  $k$ , consider the sequence of evenly spaced nodes  $v_j = I(j, k)_2$ . Now in forming the neighbourhood of  $v_j$ , we lose an interval of length (almost exactly)  $\lfloor j(2w + 1)/k \rfloor$  from the neighbourhood of  $u$ , which by Lemma 4.11 we can expect to have a proportion of  $\alpha$  nodes very close to  $\theta$ . We also gain an interval of the same length from outside  $\mathcal{N}(u)$ , which we can expect to have a proportion of  $\alpha$  nodes very close to  $\frac{1}{2}$ . This means a bias for  $v_j$  close to  $\rho \frac{k-j}{k}$ .

The following definition will allow us to express this more formally:

**Definition 4.12.** Suppose that  $\text{Hb}^*(u)$  holds. Let  $I_1 = [u - (2w + 1), u]$  and  $I_2 = [u, u + (2w + 1)]$ . Let  $|\Theta(\mathcal{N}(u))| = \rho$  and let  $\theta$  be the proportion of the nodes in  $\mathcal{N}(u)$  which are of type  $\alpha$ . Suppose that  $k \geq 1$  is even and  $\epsilon' > 0$ . For  $1 \leq j \leq k$  let  $v_j = I_2(j : k)_2$  and let  $v_{-j} = I_1(j : k)_1^-$ . We say that  $\text{Smooth}_{k, \epsilon'}(u)$  holds if both:

- For every  $j$  with  $1 \leq |j| \leq k$ ,  $|\Theta(\mathcal{N}(v_j)) - \rho \frac{k-j}{k}|/w < \epsilon'$ .
- For every  $j$  with  $1 \leq j \leq k/2$  the proportion of the nodes in  $I_1(j : k)^-$  which are of type  $\alpha$  lies in the interval  $[\theta - \epsilon', \theta + \epsilon']$ , and similarly for  $I_2(j, k)$ .

We say that  $\text{Smooth}_{k, \epsilon'}^*(u)$  holds if  $\text{Smooth}_{k, \epsilon'}(u)$  holds in the initial configuration. Figure 4 illustrates the picture for  $k = 2$ .

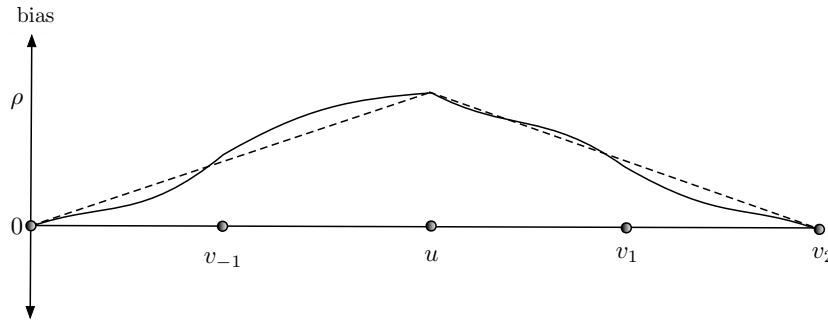


FIGURE 4. Smooth bias change for  $k = 2$ .

**Corollary 4.13** (Smoothness Corollary). *Suppose we are given that  $\text{Bb}^*(u)$  holds. For all  $k \geq 1$  and  $\epsilon' > 0$ , and for all sufficiently large  $w$ ,  $\text{Smooth}_{k,\epsilon'}^*(u)$  holds with probability  $> 1 - \epsilon'$ .*

**Proof.** Let  $u_{-1} = u - (2w + 1)$  and  $u_1 = u + (2w + 1)$ . The fact that  $u$  has borderline bias has no impact on the distributions for  $\mathcal{N}(u_{-1})$  and  $\mathcal{N}(u_1)$ . Let  $I_3 = \mathcal{N}(u_{-1})$  and  $I_4 = \mathcal{N}(u_1)$ . By the law of large numbers, for large  $w$  we can expect the proportion of  $\alpha$  nodes in each  $I_3(j : k)$  and  $I_4(j : k)$  to be close to  $\frac{1}{2}$ . The result then follows from Lemma 4.11.  $\square$

While Smoothness Corollary 4.13 tells us what can be expected from the neighbourhood of a node chosen uniformly at random from among those with borderline bias, this does not immediately allow us to infer anything about what can be expected from the neighbourhoods of each  $l_i$  and  $r_i$ . What we need is that, if we choose a node  $u$  uniformly at random and then move left (or right) until we find a first node  $v$  with high bias, then with high probability  $\text{Smooth}_{k,\epsilon'}^*(v)$  holds. Ideally, we would like to be able to apply Lemma 3.1, but that would involve establishing an appropriate form of condition (ii) from the statement of the lemma. In order to work around this condition, it turns out that what we need is a bound on the number of nodes of borderline bias which can be expected in the neighbourhood of  $u$  such that  $\text{Bb}^*(u)$  holds. The following lemma is a step in this direction:

**Definition 4.14.** *Given  $k \geq 1$ , let  $\mathcal{N}_k(u)$  be the interval  $[u - \lceil w/k \rceil, u + \lceil w/k \rceil]$ .  $\text{Bb}^*(u, k, z)$  holds if  $\text{Bb}^*(u)$  holds and there are at most  $z$  many nodes with borderline bias in  $\mathcal{N}_k(u)$  in the initial configuration.*

**Lemma 4.15.** *Suppose we are given that  $\text{Bb}^*(u)$  holds. For any  $\epsilon' > 0$  there exists  $z$  such that for  $0 \ll k \ll w$ ,  $\text{Bb}^*(u, k, z)$  holds with probability  $> 1 - \epsilon'$ .*

**Proof.** We consider the case that  $u$  has positive bias  $\rho$ , and let  $\theta$  be the proportion of the nodes in  $\mathcal{N}(u)$  which are of type  $\alpha$  (so  $\rho = (2\theta - 1)(2w + 1)$ ). The case that  $u$  has negative bias is almost identical.

First of all, we want to show that for sufficiently large  $z$ , if we step  $\lfloor z/2 \rfloor$  many nodes to the left (or right) of  $u$ , then we will very probably have bias which is well below  $\rho$ . The argument here is very similar to the proof of Corollary 4.13 – in forming the neighbourhood of  $v = u - \lfloor z/2 \rfloor$  we lose  $\lfloor z/2 \rfloor$  many nodes from  $\mathcal{N}(u)$ , with the proportion of  $\alpha$  nodes being close to  $\theta$ , and we gain the same number of new nodes, with the proportion of  $\alpha$  nodes here being close to  $\frac{1}{2}$ . Arguing more precisely, from amongst the nodes that we lose from  $\mathcal{N}(u)$ , the expected number of  $\alpha$  nodes,  $x_0$  say, is  $\theta \lfloor z/2 \rfloor$ . By applying Chebyshev's Inequality just as in the proof of Lemma 4.11, we conclude that for any  $\epsilon'' > 0$  and for all sufficiently large  $z$ ,  $\mathbf{P}(|(x_0/\lfloor z/2 \rfloor) - \theta| > \epsilon'') \ll \epsilon'$ . Now consider  $x_1$ , the number of  $\alpha$  nodes in  $\mathcal{N}(v) \setminus \mathcal{N}(u)$ . The law of large numbers tells us that for any  $\epsilon'' > 0$  and for all sufficiently large  $z$ ,  $\mathbf{P}(|(x_1/\lfloor z/2 \rfloor) - \frac{1}{2}| > \epsilon'') \ll \epsilon'$ . Combining these facts gives that for any  $m > 0$  and for all sufficiently large  $z$ :

$$\mathbf{P}(\rho - \Theta(\mathcal{N}(v)) < m) \ll \epsilon'.$$

So far then, we have considered moving  $\lfloor z/2 \rfloor$  many nodes to the left of  $u$ , and have concluded that the bias at this node  $v$  will very probably be well below  $\rho$  (a similar argument also applies, of course, moving to the right). Now we have to show that as we move left from  $v$ , so long as we remain within  $\mathcal{N}_k(u)$  the bias will very probably remain below  $\rho$ . In order to do this, we approximate the bias as we move to the left, by a biased random walk. Recall that for a biased random walk starting at value  $-m \leq -1$ , with probability  $p > \frac{1}{2}$  of going down at each step and probability  $1 - p$  of going up, the probability of ever hitting 0 is:

$$\left( \frac{1-p}{p} \right)^m.$$

So let us briefly adopt the approximation that nodes in  $\mathcal{N}(u)$  are i.i.d. random variables, each of which has probability  $\theta$  of being of type  $\alpha$ . Then, as we move left one position from a location in  $\mathcal{N}_k(u)$  to the left of  $u$ , the probability that the bias increases by 2 is  $\frac{1}{2}(1 - \theta)$ , the probability that the bias remains the same is  $\frac{1}{2}$ , and the probability that the bias decreases by 2 is  $\frac{1}{2}\theta$ . Removing those steps at which the bias does not change, we get a biased random walk with probability  $\theta$  of going down at each step and probability  $1 - \theta$  of going up. Choose  $\theta'$  with  $\frac{1}{2} < \theta' < \theta$ . Now, dropping the false assumption of independence, by taking  $k$  sufficiently large we ensure that as we take successive steps left from  $v$  inside the interval  $\mathcal{N}_k(u)$ , at each step, no matter what has occurred at previous steps, the probability of the bias increasing is less than  $\frac{1}{2}(1 - \theta')$ , the probability of the bias remaining the same is  $\frac{1}{2}$ , and the probability of the bias decreasing is greater than  $\frac{1}{2}\theta'$ . Thus, if the bias at  $u - \lfloor z/2 \rfloor$  is  $\leq \rho - m$ , then the probability that any nodes in the interval  $[u - \lceil w/k \rceil, u - \lfloor z/2 \rfloor]$  have high bias is less than  $(\frac{1-\theta'}{\theta'})^m$ .

Finally, let  $m$  be such that  $(\frac{1-\theta'}{\theta'})^m \ll \epsilon'$ , and let  $z$  be sufficiently large that, for  $v = u - \lfloor z/2 \rfloor$ ,  $\mathbf{P}(\rho - \Theta(\mathcal{N}(v)) < m) \ll \epsilon'$ .  $\square$

The reader might observe that the proof of Lemma 4.15 actually establishes something stronger, and perhaps also more natural, than claimed. Suppose we are given that  $\mathbf{Bb}^*(u)$  holds. The proof suffices to show that for any  $\epsilon' > 0$  there exists  $z$  such that, for  $0 \ll k \ll w$ , the probability there are more than  $z$  nodes of high bias in  $\mathcal{N}_k(u)$  is less than  $\epsilon'$ . One might reasonably wonder why we did not define  $\mathbf{Bb}^*(u, k, z)$  to reflect this stronger condition – that there are at most  $z$  many nodes of *high* bias (rather than borderline bias) in  $\mathcal{N}_k(u)$ . The reason is that in later counting arguments we shall need failure of the condition (in specific circumstances) to guarantee that there are, in fact, at least  $z$  many nodes of borderline bias in the relevant neighbourhood for which the condition fails.

Of course, Lemma 4.15, only restricts the number of nodes of borderline bias that will normally occur in the interval  $\mathcal{N}_k(u)$ . We have to be able to deal with a larger interval:

**Definition 4.16.** *We say that  $\text{Bb}^*(u, z)$  holds if  $\text{Bb}^*(u)$  holds and there are at most  $z$  many nodes of borderline bias in the interval  $[u - (2w + 1), u + (2w + 1)]$  in the initial configuration.*

Now note that if  $k_2$  is sufficiently large compared to  $k_1$ , if  $\epsilon'$  is sufficiently small, and if  $\text{Bb}^*(u, k_1, z)$  and  $\text{Smooth}_{k_2, \epsilon'}^*(u)$  both hold, then in the initial configuration there are at most  $z$  many nodes of borderline bias in the interval  $[u - (2w + 1), u + (2w + 1)]$ . Applying Lemmas 4.15 and 4.13 we therefore have:

**Corollary 4.17** (Few nodes of borderline bias). *Suppose we are given that  $\text{Bb}^*(u)$  holds. For any  $\epsilon' > 0$ ,  $z$  may be chosen so that for all sufficiently large  $w$ ,  $\text{Bb}^*(u, z)$  holds with probability  $> 1 - \epsilon'$ .*

We are now finally ready to prove that, with high probability when  $0 \ll w \ll n$ , the neighbourhoods of each  $l_i$  and  $r_i$  will satisfy the required smoothness conditions. We will later be able to use this in order to prove that each  $l_i$  and  $r_i$  very probably either originates a firewall or else subsides.

**Lemma 4.18** (Smoothness for  $l_i$  and  $r_i$ ). *For any node  $u$ , let  $x_u$  be the first node to the left of  $u$  which has high bias in the initial configuration. For any  $\epsilon' > 0$  and  $k \geq 1$ , if  $0 \ll w \ll n$  and  $u$  is chosen uniformly at random, then  $x_u$  is defined and  $\text{Smooth}^*(x_u, k, \epsilon')$  holds with probability  $> 1 - \epsilon'$ .*

*An analogous result holds when ‘left’ is replaced by ‘right’.*

**Proof.** Applying Corollary 4.17, choose  $z$  such that, for sufficiently large  $w$ , if we are given that  $v$  has borderline bias, then  $\text{Bb}^*(v, z)$  fails to hold with probability  $\ll \epsilon'$ , i.e. putting  $\epsilon_1 = \mathbf{P}(\neg \text{Bb}^*(v, z) | \text{Bb}^*(v))$ , choose  $z$  so that  $\epsilon_1 \ll \epsilon'$  for sufficiently large  $w$ . Let  $\epsilon_2 = \mathbf{P}(\neg \text{Smooth}_{k, \epsilon'}^*(v) | \text{Bb}^*(v))$ . Then Corollary 4.13 gives that for all sufficiently large  $w$ ,  $\epsilon_2 \ll \epsilon'/z$  – just apply the statement of the corollary to  $k$  and  $\epsilon'' \ll \epsilon'/z$ . Now define

$$\mu := \frac{\mathbf{P}(\text{Smooth}_{k, \epsilon'}^*(v) | \text{Bb}^*(v, z))}{\mathbf{P}(\neg \text{Smooth}_{k, \epsilon'}^*(v) | \text{Bb}^*(v, z))}.$$

Then

$$\mu = \frac{\mathbf{P}(\text{Smooth}_{k, \epsilon'}^*(v) \wedge \text{Bb}^*(v, z) | \text{Bb}^*(v))}{\mathbf{P}(\neg \text{Smooth}_{k, \epsilon'}^*(v) \wedge \text{Bb}^*(v, z) | \text{Bb}^*(v))} \geq \frac{1 - \epsilon_1 - \epsilon_2}{\epsilon_2} \gg \frac{z}{\epsilon'}.$$

Thus for sufficiently large  $w$ , with probability approaching 1 as  $n \rightarrow \infty$ , the ratio amongst the nodes  $v$  such that  $\text{Bb}^*(v, z)$  holds, between the number such that  $\text{Smooth}_{k, \epsilon'}^*(v)$  holds and the number such that it does not, is much greater than  $z/\epsilon'$ .

Consider the initial configuration. We define an iteration which assigns colours to nodes as follows.

Step 0. Pick a node  $t_0$  uniformly at random.



Step  $s + 1$ . Let  $v_s$  be the first node  $v$  to the left of  $t_s$  such that  $v = x_{t_s}$  or such that  $s > 0$  and  $v = t_0$ . Carry out the instructions for the first case below which applies ( $\neg$  denotes negation):

- (1) If there exists no such  $v$  then terminate the iteration, and declare that it has ‘ended prematurely’.
- (2) If  $v_s = t_0$  and  $s > 0$  then make  $t_s$  undefined and terminate the iteration.
- (3) If  $|v_s - t_s| < 2w + 1$  then colour  $t_s$  pink.
- (4) If  $\text{Bb}^*(v_s, z)$  and  $\text{Smooth}_{k, \epsilon'}^*(v_s)$  both occur, then give  $t_s$  the colours white and blue.
- (5) If  $\text{Bb}^*(v_s, z)$  and  $\neg \text{Smooth}_{k, \epsilon'}^*(v_s)$  both occur, then give  $t_s$  the colours white and red.
- (6) If  $v_s$  satisfies  $\neg \text{Bb}^*(v_s, z)$ , then give  $t_s$  the colour black.

In cases (3)–(6), define  $t_{s+1} = v_s - (2w + 1)$ , unless  $t_0$  lies in the interval  $[v_s - (2w + 1), v_s]$ , in which case terminate the iteration.

This completes the description of the iteration.

First note that the probability that the iteration terminates prematurely can be made arbitrarily small by taking  $n$  large, and similarly that we may assume there are  $t_s$  of all colours. Now let  $S$  be the greatest  $s$  such that  $t_s$  is defined when the iteration terminates, and let  $\pi$  be the proportion of the  $t_s$ ,  $s \leq S$ , such that  $t_s$  is coloured black. Amongst the nodes  $u$  which have borderline bias, let  $\rho$  be the ratio between the number for which  $\text{Bb}^*(u, z)$  holds and the number for which it does not – so that for large  $n$ ,  $\rho$  can be expected to be close to  $(1 - \epsilon_1)/\epsilon_1$ . In order to find an upper bound for  $\rho$ , first let us find an upper bound for the number of nodes  $u$  such that  $\text{Bb}^*(u, z)$  holds. Let  $v_s$  be as defined in the iteration. However we colour  $t_s$ , there can be at most  $z$  many nodes  $u$  in  $I_s := (v_s - (2w + 1), v_s]$  which satisfy  $\text{Bb}^*(u, z)$ , giving an upper bound of  $(S + 1)z$ . Now, let us find a lower bound for the number of nodes with borderline bias for which  $\text{Bb}^*(u, z)$  fails. If  $t_s$  is coloured pink, then we are not guaranteed any nodes in  $I_s$  of borderline bias for which  $\text{Bb}^*(u, z)$  fails. If  $t_s$  is coloured white, then the same applies. If  $t_s$  is coloured black, however, we are guaranteed at least  $z$  many nodes of borderline bias in  $I_s$  for which  $\text{Bb}^*(u, z)$  fails. We therefore get a lower bound of  $(S + 1)\pi z$ . Thus:

$$\rho \leq \frac{z(S + 1)}{\pi(S + 1)z} = \frac{1}{\pi}.$$

We chose  $z$  previously, so that for all sufficiently large  $w$ ,  $\epsilon_1 \ll \epsilon'$ . Since  $\rho$  is close to  $(1 - \epsilon_1)/\epsilon_1$  for large  $n$ , we infer that for all sufficiently large  $w$ , with probability tending to 1 as  $n \rightarrow \infty$ , we have  $\rho \gg 1/\epsilon'$ , so that  $\pi \ll \epsilon'$ . For large  $n$ , the probability  $t_s$  is coloured pink is less than  $(2w + 1)e^{-w/d}$  (with  $d$  as in Definition 4.2). So, for sufficiently large  $w$ , with probability tending to 1 as  $n \rightarrow \infty$ , the proportion of the  $t_s$  which are not coloured white is  $\ll \epsilon'$ .

Now let  $\beta$  be the proportion of the  $t_s$  which are coloured red. Amongst the nodes such that  $\mathbf{Bb}^*(u, z)$  holds, let  $\gamma$  be the ratio between the number for which  $\mathbf{Smooth}_{k, \epsilon'}^*(u)$  holds, and the number for which it does not – so that  $\gamma$  can be expected to be close to  $\mu$  (as defined earlier) for large  $n$ . Again, let us form an upper bound for  $\gamma$ . However  $t_s$  is coloured, we get at most  $z$  many nodes in the interval  $I_s$  for which  $\mathbf{Bb}^*(u, z) \wedge \mathbf{Smooth}_{k, \epsilon'}^*(u)$  holds. If  $t_s$  is coloured red then we get at least one node in the interval  $I_s$  for which  $\mathbf{Bb}^*(u, z) \wedge \neg \mathbf{Smooth}_{k, \epsilon'}^*(u)$  holds. Thus:

$$\gamma \leq \frac{z(S+1)}{\beta(S+1)} = \frac{z}{\beta}.$$

We previously observed that, for all sufficiently large  $w$ ,  $\mu \gg z/\epsilon'$ . Thus, for all sufficiently large  $w$ , with probability tending to 1 as  $n \rightarrow \infty$ , we have that  $\gamma \gg z/\epsilon'$ , so that  $\beta \ll \epsilon'$ . Since we have already established that, for  $0 \ll w \ll n$ , the probability  $t_s$  is not coloured white is  $\ll \epsilon'$ , we conclude that for  $0 \ll w \ll n$  the probability  $t_s$  is coloured blue is  $> 1 - \epsilon'$ . Now for large  $n$ , the probability that  $t_s$  is coloured blue is less than the probability, for  $u$  chosen uniformly at random, that  $v_u$  is defined and  $\mathbf{Bb}^*(v_u, z) \wedge \mathbf{Smooth}_{k, \epsilon'}^*(v_u)$  holds, which concludes the proof.  $\square$

So Lemma 4.18 tells us that, for some fixed  $k_1 \geq 1$  and  $\epsilon' > 0$  (to be chosen later), we will be able to define one of the  $Q_j$  (in fact this will be  $Q_2$ ) to be the event that, when  $1 \leq i \leq k_0$  and  $u = l_i$  or  $u = r_i$ ,  $\mathbf{Smooth}_{k_1, \epsilon'}^*(u)$  holds. This event will then occur with probability tending to 1 as  $w \rightarrow \infty$ .

Shortly we shall go on to analyse what can be expected to occur in the interval  $[l_{k_0}, r_{k_0}]$  once the process of type changing begins. First of all, however, we have to establish one more thing for the initial configuration. Although we are working for fixed  $\tau$ , in order to be able to talk about intervals which are ‘close’ to being stable, it is also useful to be able to talk about intervals which are stable with respect to  $\tau' < \tau$ :

**Definition 4.19.** *We say that an interval of length  $w + 1$  is  $\tau'$ -stable if it contains  $\tau'(2w + 1)$  many  $\alpha$ -nodes or  $\tau'(2w + 1)$  many  $\beta$  nodes.*

**Lemma 4.20** (Nothing close to stable in the vicinity of  $u_0$ ). *Suppose  $\kappa < \tau' < \tau$ . For any  $u$ , let  $x_u$  be the first node to the left of  $u$  which, in the initial configuration, either has high bias or else belongs to an interval which is  $\tau'$ -stable. For any  $\epsilon' > 0$ , if  $0 \ll w \ll n$  then the following occurs with probability  $> 1 - \epsilon'$  for  $u$  chosen uniformly at random:  $x_u$  is defined and no node in  $[x_u - 2w, x_u]$  belongs to a  $\tau'$ -stable interval.*

*An analogous result holds when ‘left’ is replaced by ‘right’.*

**Proof.** For a moment, consider working with  $\tau'$  rather than  $\tau$ . Recall that, in Section 3 we showed that for constants  $C_1$  and  $C_2$ :

$$C_1 \cdot \left( \frac{(\frac{1}{2} - \tau')^{1-2\tau'}}{(1 - \tau')^{2(1-\tau')}} \right)^w \leq \frac{P_{\text{unhap}}}{P_{\text{stab}}} \leq C_2 \cdot \left( \frac{(\frac{1}{2} - \tau')^{1-2\tau'}}{(1 - \tau')^{2(1-\tau')}} \right)^w$$

Thus, for any  $k$  and for sufficiently large  $w$ , the probability that a randomly chosen node is unhappy is more than  $kw$  times the probability that a randomly chosen node belongs to a  $\tau'$ -stable interval. Now, since  $\tau > \tau'$ , this only makes unhappy nodes *more* likely in the initial configuration. Thus, working with  $\tau$ , for any  $k$  and for all sufficiently large  $w$ , the probability that a randomly chosen node is unhappy is more than  $kw$  times the probability that a randomly chosen node belongs to a  $\tau'$ -stable interval. The result then follows from Lemma 3.1.  $\square$

Although we have already established a lot concerning what can be expected from the initial configuration, so far we have only gotten as far as precisely defining the first event  $Q_0$ . Here are our second and third events:

**Definition 4.21** (Defining  $Q_1$ ). *Once and for all, fix some  $\tau_0$  with  $\kappa < \tau_0 < \tau$ . Let  $Q_1$  be the event that  $l_i$  and  $r_i$  are defined for all  $1 \leq i \leq k_0$ , and there do not exist any  $\tau_0$ -stable subintervals of  $[l_{k_0}, r_{k_0}]$  in the initial configuration.*

Note that, by Lemma 4.20, when  $k_0 \ll w$  the probability that  $Q_1$  does not occur is  $\ll \epsilon$ .

**Definition 4.22** (Defining  $Q_2$ ). *Let  $\tau_0$  be as in Definition 4.21. Once and for all, choose  $k_1$  such that  $\frac{1}{k_1} \ll \tau - \tau_0$ , and choose  $k_2$  and  $\epsilon_0$  such that  $k_1 \ll k_2 \ll \frac{1}{\epsilon_0}$  and  $k_2$  is a multiple of  $k_1$ . We define  $Q_2$  to be the event that all the  $l_i$  and  $r_i$  are defined for  $1 \leq i \leq k_0$ , and that when  $u = l_i$  or  $u = r_i$ ,  $\text{Smooth}_{k_2, \epsilon_0}^*(u)$  holds.*

By Lemma 4.18, when  $k_0 \ll w$  the probability that  $Q_2$  does not occur is  $\ll \epsilon$ .

**The process to completion.** Having established a clearer picture of what can be expected from the initial configuration, we now look to understand what will happen in the early stages, in the neighbourhood of each  $l_i$  or  $r_i$ . First of all, we must establish that these nodes can be expected to complete.

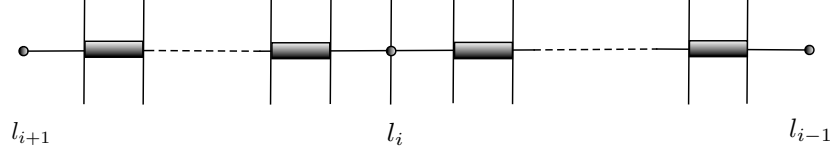
**Lemma 4.23** ( $l_i$  and  $r_i$  complete). *For any  $\epsilon' > 0$ , if  $0 \ll w \ll n$  then for all  $i \in [1, k_0]$ ,  $l_i$  and  $r_i$  will (be defined and will) complete with probability  $> 1 - \epsilon'$ .*

**Proof.** We form an upper bound for the probability that  $l_i$  will fail to complete (the proof for  $r_i$  is essentially identical). Suppose that  $Q_0$  holds, and that  $w$  is large. As before, it is convenient to define  $l_0 = u_0$ . Fix  $i < k_0$  and let  $I_1 = [l_{i+1}, l_i]$  and  $I_2 = [l_i, l_{i-1}]$ . Let  $k$  be the greatest such that, when  $1 \leq j \leq k$ ,  $I_1(j : k)$  and  $I_2(j : k)$  are of length  $\geq w + 1$ . For  $1 \leq j \leq \lfloor k/2 \rfloor$  define:

$$J_j := I_1(j : k) \cup I_1(j : k)^- \cup I_2(j : k) \cup I_2(j : k)^-.$$

Figure 5 below shows, as an example, what  $J_2$  looks like:

For  $1 \leq j \leq \lfloor k/2 \rfloor$ , let  $P_j$  be the event that a node in  $J_j$  changes type, and note that  $P_3$  cannot occur until  $P_2$  has occurred,  $P_4$  cannot occur until  $P_3$  has occurred, and so

FIGURE 5.  $J_2 = I_1(2 : k) \cup I_1(2 : k)^- \cup I_2(2 : k) \cup I_2(2 : k)^-$ .

on. Now the basic idea is that if completion fails to occur, then the sequence of events  $P_2, \dots, P_{\lfloor k/2 \rfloor}$  must occur before any stage at which there are no unhappy nodes in the neighbourhood of  $l_i$ .

We label certain stages as being a ‘step towards completion’, and certain others as being a ‘step towards failure of completion’ (while some stages are labelled as neither). These labels do not fully represent all aspects of the process, but contain enough information for our purposes.

**Steps towards completion.** So long as there are unhappy nodes in the neighbourhood of  $l_i$ , we label any stage at which a node in this neighbourhood swaps type as a *step towards completion*. Subsequent to any stage at which  $l_i$  completes, we label every stage as a step towards completion.

**Steps towards failure of completion.** If  $1 \leq j < \lfloor k/2 \rfloor$  is the greatest such that  $P_j$  has occurred prior to stage  $s$  or no  $P_j$  has occurred and  $j = 1$ , and if  $P_{j+1}$  occurs at stage  $s$ , then we label  $s$  a *step towards failure of completion*.

Now at any stage  $s$  at which some  $P_j$  for  $j \leq \lfloor k/2 \rfloor$  is yet to occur, and at which there are unhappy nodes in the neighbourhood of  $l_i$ , the probability of  $s$  being a step towards failure of completion, is at most  $4(w+1)$  times the probability of it being a step towards completion (since there are at most  $4(w+1)$  times as many nodes which, if chosen to swap, will cause a step towards failure of completion, as those which will cause a step towards completion). Choosing  $d' > d$  we get that, since  $Q_0$  holds, for all sufficiently large  $w$ ,  $\lfloor k/2 \rfloor > e^{w/d'}$ . We may therefore consider the first  $e^{w/d'}$  many stages which are steps either towards completion or failure of completion<sup>4</sup> and, for large  $w$ , consider the probability that at most  $2w$  of these are steps towards completion. By the law of large numbers, this probability tends to 0 as  $w \rightarrow \infty$ . Now if there do not exist unhappy nodes of both types in the interval  $[l_i - 2w, l_i]$  in the initial configuration (which is the case with probability tending to 1 as  $w \rightarrow \infty$ , since it holds if  $\text{Smooth}^*(l_i, k', \epsilon'')$  holds for large  $k'$  and small  $\epsilon''$ ), then  $2w + 1$  many steps towards completion prior to  $P_{\lfloor k/2 \rfloor}$  occurring, suffices to ensure completion for  $l_i$ .  $\square$

**Definition 4.24** (Defining  $Q_3$ ). *We define  $Q_3$  to be the event that all the  $l_i$  and  $r_i$  (for  $1 \leq i < k_0$ ) are defined and complete.*

<sup>4</sup>In order to ensure the existence of  $e^{w/d'}$  many such stages it is momentarily convenient to adopt the convention that the process continues after there are no unhappy nodes  $u \in [1, n]$ , but with nothing occurring at such stages, and that such stages are also labelled steps towards completion.

**Lemma 4.25** (The required dichotomy). *Suppose that  $w$  is large and that  $Q_j$  holds for all  $j \leq 3$ . Then for  $i < k_0$ ,  $l_i$  and  $r_i$  will each either subside or originate a firewall.*

**Proof.** We prove the result for  $l_i$ , and the proof for  $r_i$  is essentially identical.

Note first that the choice of  $k_1$  in Definition 4.22 means, in particular, that  $10w/k_1$  type changes in any given neighbourhood cannot create stable intervals (given that  $Q_1$  holds). Note also, that satisfaction of  $Q_2$  suffices to ensure that there are not unhappy nodes of both types in the neighbourhood of  $l_i$  in the initial configuration. Now suppose that  $l_i$  completes at stage  $s$  and has positive bias in the initial configuration (the case for negative bias is essentially identical).

It is useful at this point to establish names for a number of relevant intervals. We let  $u_1 = l_i - (2w + 1)$  and  $u_2 = l_i + (2w + 1)$ . Then we define:

- $J = \mathcal{N}(u_1) \cup \mathcal{N}(l_i) \cup \mathcal{N}(u_2)$ .
- $I = [u_1, u_2]$ .
- $I_1 = [u_1, l_i]$ ,  $I_2 = [l_i, u_2]$ .
- $K_j^1 = I_1(i : k_1)^- \cup I_2(j : k_1)$ .
- $K_j^2 = I_1(i : k_2)^- \cup I_2(j : k_2)$ .

In Definition 4.22 we assumed that  $k_2$  is a multiple of  $k_1$ , so we may let  $m$  be such that  $k_2 = mk_1$ . It is convenient to assume that  $k_2$  is even. Now we divide into two cases.

**$l_i$  subsides.** First of all, suppose that at stage  $s$ , there is a  $\beta$ -node  $u$  in the interval  $K_1^1$ . Then  $u$  must be happy at stage  $s$ . The fact that  $Q_2$  is satisfied, together with the fact that  $l_i$  completes at stage  $s$ , means that prior to stage  $s$ , the only type changes in the interval  $J$  are from type  $\beta$  to type  $\alpha$ , so that  $u$  must be happy at *every* stage  $\leq s$ . Now, since  $k_1 \ll k_2 \ll \frac{1}{\epsilon_0}$  and  $\text{Smooth}_{k_2, \epsilon_0}^*(l_i)$  holds, any nodes in  $I - (K_1^1 \cup K_2^1)$  have lower bias than  $u$ , and hence are happy, in the initial configuration. It then follows by induction on the stages  $\leq s$  that no node in  $J - (K_1^1 \cup K_2^1)$  changes type prior to stage  $s$ . In order to see this suppose that it holds prior to stage  $s' \leq s$ . Then at stage  $s'$ , if  $v \in I - (K_1^1 \cup K_2^1)$ , it still has lower bias than  $u$  and so cannot change type from  $\beta$  to  $\alpha$ , since  $u$  is happy so  $v$  must be. If  $v \in J - I$ , then  $v$  changing type would contradict the fact that  $l_i$  completes.

We therefore get at most  $|(K_1^1 \cup K_2^1)| < 10w/k_1$  many type changes in the interval  $J$  prior to completion. As observed above, this means that no stable intervals are created and  $l_i$  subsides, as required.

**$l_i$  originates a firewall.** So suppose instead that, at stage  $s$ , all nodes in the interval  $K_1^1$  are of type  $\alpha$ . Given  $m$  as above, another way of putting this, is that all nodes in  $\bigcup_{j \leq m} K_j^2$  are of type  $\alpha$  at stage  $s$ . We now show by induction on  $r \geq m$  that, when  $r \leq k_2/2$ , any nodes in  $K_r^2$  must be of type  $\alpha$  at stage  $s$  – i.e. that all nodes in  $\mathcal{N}(l_i)$  are  $\alpha$  nodes at stage  $s$ . So suppose that  $m \leq r < k_2/2$  and that the hypothesis holds for all  $r' \leq r$ . Consider  $u \in K_{r+1}^2$ . Let  $\rho$  be the bias of  $l_i$  in the initial configuration. First let us form a lower bound for the bias of  $u$  in the initial configuration. The fact that  $Q_2$  holds means that the leftmost and rightmost nodes in  $K_r^2$  have bias at least  $\rho - \frac{r}{k_2}\rho - \epsilon_0 w$ . Then, since

the bias can change by at most 2 if we move left or right one node, we conclude that in the initial configuration  $u$  has bias

$$\rho_1 \geq \rho - \left( \frac{r}{k_2} \rho + \epsilon_0 w + \frac{2(2w+1)}{k_2} \right).$$

Now we have to take into account all of the  $\beta$  nodes in  $\bigcup_{j \leq r} K_r^2$  which have changed type. In fact, so that we can be sure that each change of type affects the bias of  $u$ , we shall consider just those which lie between  $l_i$  and  $u$ . Let  $\theta$  be the proportion of the nodes in  $\mathcal{N}(l_i)$  which are  $\alpha$  nodes – so that  $\rho = (2\theta - 1)(2w + 1)$ . Then the number of  $\beta$  nodes in  $\bigcup_{j \leq r} K_j^2$  in the initial configuration, which lie between  $l_i$  and  $u$ , is at least:

$$(1 - \theta - \epsilon_0)(2w + 1)r/k_2.$$

Each change of type for one of these nodes means an increase of 2 in the bias of  $u$ , so that at stage  $s$ ,  $u$  has bias:

$$\begin{aligned} \rho_2 &\geq \rho - \left( \frac{r}{k_2} \rho + \epsilon_0 w + \frac{2(2w+1)}{k_2} \right) + (1 - \theta - \epsilon_0)2(2w + 1)r/k_2. \\ &= \rho + (2w + 1) \left( \frac{(1 - \theta)2r}{k_2} - \frac{\epsilon_0 w}{2w + 1} - \frac{2}{k_2} - \frac{r}{k_2}(2\theta - 1) - \frac{\epsilon_0 2r}{k_2} \right) \end{aligned}$$

We are left to compare the terms

$$\frac{(1 - \theta)2r}{k_2}, \quad \frac{\epsilon_0 w}{2w + 1}, \quad \frac{2}{k_2}, \quad \frac{r}{k_2}(2\theta - 1) \text{ and } \frac{\epsilon_0 2r}{k_2}.$$

Since  $1/k_2 \gg \epsilon_0$ , the second term is much smaller than the third. Since  $(1 - \theta) \in (0.5, 0.65)$ ,  $r/k_2 \geq 1/k_1$  and  $k_1 \ll k_2$ , the first term is much larger than the third. Since  $\epsilon_0$  is small, the first term is also much larger than the last. The result then follows for large  $w$ , since  $2 - 2\theta$  is always more than double  $2\theta - 1$  for  $\kappa < \tau < \frac{1}{2}$ , meaning that the first term is more than double the fourth term, and thus  $\rho_2 \geq \rho$ , meaning that if  $u$  is a  $\beta$  node, it will be unhappy.  $\square$

**Lemma 4.26** (Reasonable chance of firewall). *There exists  $\delta > 0$  which does not depend on  $w$ , such that if  $1 \leq i < k_0$ , then  $l_i$  originates a firewall with probability  $> \delta$  (and similarly for  $r_i$ ).*

**Proof.** Let  $\theta$  be the proportion of the nodes in  $\mathcal{N}(l_i)$  which are of type  $\alpha$ . Suppose that  $l_i$  has positive bias  $\rho$ . Just as we did in the proof of Lemma 4.15, let us initially adopt the approximation that the nodes in  $\mathcal{N}(l_i)$  are i.i.d. with probability  $\theta$  of being an  $\alpha$  node. Let  $K_1^1$  be as defined in the proof of Lemma 4.25. From the argument given there, it suffices to show that there exists  $\delta > 0$  which does not depend on  $w$ , such that the probability of  $l_i$  completing with all nodes in  $K_1^1$  being of type  $\alpha$ , is greater than  $\delta$ .

We consider a process, consisting of  $\ell := \lceil \frac{w}{k_1} \rceil + 1$  many steps.

**Step 0.** If  $l_i$  is of type  $\beta$  then define  $\rho_0 = \rho + 2$ , and otherwise define  $\rho_0 = \rho$ .

**Step  $s > 0$**  (with  $s \leq \ell$ ). Consider the two nodes  $l_i - s$  and  $l_i + s$ . Let  $x_s$  be the number of these which are  $\beta$  nodes (so  $x_s \in \{0, 1, 2\}$ ). Let  $y_s$  be  $\Theta^*(\mathcal{N}(l_i - s)) - \Theta^*(\mathcal{N}(l_i - s + 1))$ . Then define  $\rho_s = \rho_{s-1} + 2x_s + y_s$ .

So, given the approximation that the nodes in  $\mathcal{N}(l_i)$  are i.i.d. with probability  $\theta$  of being an  $\alpha$  node, this gives a biased random walk. If  $\rho_s > \rho$  for all  $s$  then, given that  $l_i$  completes, all the nodes to the left of  $l_i$  in  $K_1^1$  must be  $\alpha$  nodes at completion. We may also consider the mirror image process, in which  $y_s$  is defined in terms of the bias at  $l_i + s$  and  $l_i + s - 1$  instead of the bias at  $l_i - s$  and  $l_i - s + 1$ . Suppose that this process gives a set of values  $\rho'_s$ . Again, if  $\rho'_s > \rho$  for all  $s$  then, given that  $l_i$  completes, all the nodes to the right of  $l_i$  in  $K_1^1$  must be  $\alpha$  nodes at completion. It therefore suffices to show that  $\rho_s - \rho_{s-1}$  is more likely to be positive than negative, for  $0 < s \leq \ell$ .

The probability that  $\rho_s - \rho_{s-1} = -2$  is  $\frac{1}{2}\theta^3$ . The probability that  $\rho_s - \rho_{s-1} = 0$  is  $\frac{1}{2}\theta \cdot 2\theta(1 - \theta) + \frac{1}{2}\theta^2 = \theta^2(1 - \theta) + \frac{1}{2}\theta^2$ . Therefore the probability that  $\rho_s - \rho_{s-1} \in \{2, 4, 6\} = 1 - \frac{1}{2}\theta^3 - \theta^2(1 - \theta) - \frac{1}{2}\theta^2 > \frac{1}{2}$ .

Finally, we drop the false assumption of independence, in the same way as in the proof of Lemma 4.15. In the random walk described above, let  $p > \frac{1}{2}$  be the probability that  $\rho_s - \rho_{s-1} > 0$ . Now choose  $p'$  with  $\frac{1}{2} < p' < p$ . For sufficiently large  $k_1$ , when we drop the assumption of independence, the actual probability that  $\rho_s - \rho_{s-1} > 0$  at each step is greater than  $p'$ , no matter what has occurred at previous steps.  $\square$

**Definition 4.27** (Defining  $Q_4$  and choosing  $k_0$ ). *We let  $Q_4$  be the event that one of the  $l_i$   $i < k_0$  is defined and originates a firewall, and that the same holds for some  $r_j$ ,  $j < k_0$ . According to Lemma 4.26, given  $\epsilon > 0$  we can choose  $k_0$  once and for all, which is large enough such that the probability  $Q_4$  does not occur is  $\ll \epsilon$  for  $0 \ll w \ll n$ .*

So far then, we have proved that, for any  $\epsilon' > 0$ , for  $0 \ll w \ll n$ , with probability  $> 1 - \epsilon'$  there will be a least  $i$  such that  $l_i$  is defined and originates a firewall, there will be a least  $j$  such that  $r_j$  is defined and originates a firewall, and both  $l_i$  and  $r_j$  will be at a distance  $> e^{w/d}$  from  $u_0$ . It remains to prove that the firewalls originated at  $l_i$  and  $r_i$  will spread until  $u_0$  is contained in one of them. The following lemma therefore completes the proof for the simple model.

**Lemma 4.28** ( $u_0$  ultimately joins a firewall). *Suppose that all  $Q_i$  hold for  $i \leq 4$ . Let  $i$  be the least such that  $l_i$  is defined and originates a firewall, and let  $j$  be least such that  $r_j$  is defined and originates a firewall. For any  $\epsilon' > 0$ , for all sufficiently large  $w$ , with probability  $> 1 - \epsilon'$ ,  $u_0$  will eventually be contained in one of the two firewalls originated at  $l_i$  and  $r_i$ .*

**Proof.** If the firewalls originated at  $l_i$  and  $r_j$  are of the same type, then the result is immediate. So suppose that the firewall originated at  $l_i$  is of type  $\alpha$  and the firewall originated at  $r_j$  is of type  $\beta$ .

As we describe the argument we initially state two facts ( $\dagger_1$ ) and ( $\dagger_2$ ) without proof. Once the outline of the argument is complete, we then provide proofs for these facts.



First, note that there are certain type changes within the interval  $[l_i, r_j]$  which we are not presently concerned with. If  $u$  is in the neighbourhood of some  $x$  which is  $l_{i'}$  for  $1 \leq i' < i$  or  $r_{j'}$  for  $1 \leq j' < j$ , and changes type at a stage which is less than or equal to that at which  $x$  completes, then we say that this change of type is *previous*. Ignoring changes which are previous, it is then easy to formalise the first stage  $s(I)$  at which either of the two firewalls originated at  $l_i$  or  $r_j$  ‘have influence’ on any given subinterval  $I$  of  $[l_i, r_j]$ :  $s(I)$  is the first stage at which any node in  $I$  has a change of type which is not previous. The fact that all  $Q_k$  for  $k \leq 4$  are satisfied, means that  $s(I)$  must be defined for any subinterval  $I$  of  $[l_i, r_j]$ .

For  $\gamma \in \{\alpha, \beta\}$  we also define  $s^\gamma(I)$  to be the first stage at which a node in  $I$  has a change of type to  $\gamma$  which is not previous – note that, unlike  $s(I)$ , these values may be undefined (we write  $s^\gamma(I) \downarrow$  to indicate that  $s^\gamma(I)$  is defined). Now let  $u_1 = u_0 - (2w + 1)$  and let  $u_2 = u_0 + (2w + 1)$ . Let  $I_0 = \mathcal{N}(u_0)$ ,  $I_1 = \mathcal{N}(u_1)$  and let  $I_2 = \mathcal{N}(u_2)$ .

- (†<sub>1</sub>) For any  $\epsilon' > 0$ , for all sufficiently large  $w$ , the probability that there does not exist  $\gamma \in \{\alpha, \beta\}$  such that both  $s^\gamma(I_1) \downarrow = s(I_1)$  and  $s^\gamma(I_2) \downarrow = s(I_2)$ , is  $\ll \epsilon'$ .

So suppose that there does exist such  $\gamma$ , and suppose  $\gamma = \beta$  (the case  $\gamma = \alpha$  is similar).

- (†<sub>2</sub>) Let  $v \in [l_i, r_j]$  be the rightmost node such that either  $v = l_i$  or  $s^\alpha(\mathcal{N}(v)) \downarrow < s(I_1)$ . For any  $\epsilon' > 0$ , for all sufficiently large  $w$ , the probability that  $|u_0 - v| < e^{w/d}$  is  $\ll \epsilon'$ .

Let us suppose that  $|u_0 - v| \geq e^{w/d}$ . Now consider  $I_0$ , the neighbourhood of  $u_0$ . Let  $I_3$  be the rightmost  $w$  many nodes in this neighbourhood, and let  $I_4$  be the leftmost  $w$  many nodes. Let  $\delta = 1 - 2\tau$ , so that if any node  $x$  has bias 0 in the initial configuration then  $w\delta$  many type changes in  $\mathcal{N}(x)$ , all to type  $\beta$  suffice to give  $x$  almost exactly borderline bias (less than borderline bias plus 3 to be more precise). By the Central Limit Theorem, for any  $\epsilon' > 0$ , for all sufficiently large  $w$ , the probability that any node in  $I_0$  or  $I_1$  has bias in the initial configuration which is of modulus  $> w\frac{1}{3}\delta$  is  $\ll \epsilon'$ . Suppose that this is not the case. Now no node in  $I_4$  can become unhappy until there have been  $> \frac{2}{3}\delta w$  many changes to  $\beta$  type in  $I_3$  (with room to spare). Similarly no node in  $I_1$  can become unhappy until there have been at least  $\frac{2}{3}\delta w$  many changes to  $\beta$  type in  $I_4$ . So at stage  $s(I_1)$  we conclude that there have been at least  $\frac{4}{3}\delta w$  many changes to  $\beta$  type in the neighbourhood of  $u_0$ . This means that all nodes in  $I_0$  now have bias at most  $w\frac{1}{3}\delta - w\frac{8}{3}\delta$ . Thus *all*  $\alpha$  nodes in  $I_0$  must be unhappy at stage  $s(I_1)$ . We can now define a notion of completion for  $u_0$ . We say that  $u_0$  completes at stage  $s > s(I_1)$  if:

- (1) No node in  $I_0$  is unhappy at stage  $s$ , and this is not true for any  $s' < s$  with  $s' > s(I_1)$ .
- (2) Letting  $v$  be defined as in (†<sub>2</sub>), there exists  $x_0$  with  $v + 2w < x_0 < u_0 - 2w$ , such that by the end of stage  $s$ , no node in  $[x_0 - w, x_0]$  has had a change of type which is not previous.

We can then argue, precisely as in Lemma 4.23, that  $u_0$  almost certainly completes. In this case it then follows that  $u_0$  ultimately belongs to a firewall, which includes the interval  $[u_0, r_j]$ . To complete the proof, we are therefore left to verify  $(\dagger_1)$  and  $(\dagger_2)$ .

We verify  $(\dagger_1)$  in such a way that almost precisely the same proof suffices to verify  $(\dagger_2)$  also. Given  $\epsilon' > 0$ , choose  $k \gg \frac{1}{\epsilon'}$ . For all sufficiently large  $w$ , it follows from our earlier analysis that, in fact, the probability that  $|u_0 - l_1| < (k+1)\lceil e^{w/d} \rceil$  or  $|u_0 - r_1| < (k+1)\lceil e^{w/d} \rceil$  is  $\ll \epsilon'$ . So suppose that neither of these possibilities hold. Now, for each instance  $\mathcal{Z}$  of the process on the circle of  $n$  nodes for which there does not exist  $\gamma \in \{\alpha, \beta\}$  such that both  $s^\gamma(I_1) \downarrow = s(I_1)$  and  $s^\gamma(I_2) \downarrow = s(I_2)$ , we wish to show that there are  $k$ -many distinct others which we can label as *corresponding* to  $\mathcal{Z}$ , each of which occurs with the same probability as  $\mathcal{Z}$  and for which there *does* exist  $\gamma$  of the kind described (and for which all  $Q_n$  for  $n \leq 4$  hold for the same  $u_0$ ). We also require that, for distinct  $\mathcal{Z}$  and  $\mathcal{Z}'$  for which  $\gamma$  as required does not exist, the two corresponding sets of processes have no intersection. The  $k$ -many distinct processes corresponding to  $\mathcal{Z}$  are easily defined. Given that  $\gamma$  does not exist, we must have that  $s(I_1) = s^\alpha(I_1)$  and that  $s(I_2) = s^\beta(I_2)$ . The processes corresponding to  $\mathcal{Z}$  are the ‘rotations’ of the entire process  $\mathcal{Z}$  by  $n\lceil e^{w/d} \rceil$  many nodes to the right, for  $1 \leq n \leq k$ . We therefore conclude that the probability that  $\gamma$  does not exist as required, is  $< \frac{1}{k} \ll \epsilon'$ .

As remarked above, an almost identical proof then suffices to establish  $(\dagger_2)$ .  $\square$

**The standard model.** The difficulty that arises when one moves to the standard model, is that when there are different numbers of unhappy  $\alpha$  and  $\beta$  nodes, it is no longer true that every unhappy node is equally likely to be chosen as part of a swapping pair. If there are more unhappy  $\alpha$  nodes than unhappy  $\beta$  nodes at a given stage, for example, then unhappy  $\beta$  nodes belong to more unhappy pairs of opposite type than do their unhappy  $\alpha$  counterparts, and so are more likely to be chosen as part of an unhappy pair. This potentially complicates our proof of Lemma 4.23, that each  $l_i$  and  $r_i$  will almost certainly complete, for example, since it may make ‘steps towards completion’ (as defined in that proof) less likely. The solution, just as in [2], is to use technology developed by Wormald [21] in order to show that we can sufficiently accurately model the discrete process with a continuous one which is governed by a system of differential equations, and thereby demonstrate that the number of unhappy nodes of each type actually remains very evenly balanced. The remainder of this section is very similar to the corresponding argument in [2]. We give the full proof since there are some minor differences, and to make the paper as self-contained as possible. We aim to prove the following fact, which shows that our proof for the simple model suffices for the standard model as well:

- ( $\diamond$ ) Suppose  $w, \tau$  (with  $\kappa < \tau < \frac{1}{2}$ ) and  $\epsilon > 0$  are fixed (where  $\epsilon$  is the value we have fixed throughout this section, and which played a role in the definition of  $k_0$ ). Let  $l_{k_0}$  and  $r_{k_0}$  be as defined previously. For any  $\epsilon' > 0$ , when  $n$  is sufficiently large the following holds with probability  $> 1 - \epsilon'$ : there exists a first stage at which there are no unhappy nodes in the interval  $[l_{k_0}, r_{k_0}]$  and at all stages up to this the total

number of unhappy  $\alpha$  nodes divided by the total number of unhappy  $\beta$  nodes lies in the interval  $[1 - \epsilon', 1 + \epsilon']$ .

The very basic idea is as follows. For the remainder of the section we suppose that  $w$ ,  $\tau$  and  $\epsilon$  are fixed. With high probability, for large  $n$  we will have in the initial configuration that the proportion of nodes which are unhappy and of type  $\alpha$  is roughly equal to the proportion of nodes which are unhappy and of type  $\beta$ . Briefly, however, let us make the simplification that these proportions will be *exactly* equal. We then want to show that the process can be sufficiently accurately modelled, for large  $n$ , by a system of differential equations, which are entirely symmetric in  $\alpha$  and  $\beta$ . This symmetry means that the solution to the system of differential equations must describe an evolution in which there are always precisely equal numbers of unhappy  $\alpha$  and  $\beta$  nodes. Of course, we then have to deal with the fact that the numbers of unhappy  $\alpha$  and  $\beta$  nodes in the initial configuration need not actually be exactly equal, but this turns out not to present too many problems.

As discussed in [2], there are, however, some further complications which arise immediately as one looks to apply the Wormald machinery. The method applies to a process in which the state of the system at any given moment in time is described by an  $\ell$ -dimensional vector of real numbers, where  $\ell$  remains fixed, and we then look to approximate the discrete process by a continuous one as  $n \rightarrow \infty$ . How can we describe the configuration at any given stage by an  $\ell$ -dimensional vector, where  $\ell$  is independent from  $n$ ? Let  $C_n$  be the graph which is a cycle of size  $n$ . Up until now, then, we have been considering processes unfolding on each  $C_n$ . From now on we consider also a value  $L$  which depends on  $w$ , but not on  $n$ . Generally we shall work under the assumption that  $w \ll L \ll n$ . For the sake of simplicity we assume that  $L$  divides  $n$  (but everything that follows is easily modified to deal with the possibility that this is not the case). As we consider the process unfolding on  $C_n$ , we consider also a parallel process on  $G_n$ , which is a disjoint union of cycles of length  $L$ . More precisely, nodes  $u$  and  $v$  are connected in  $G_n$  iff  $\lfloor u/L \rfloor = \lfloor v/L \rfloor$  and  $u \equiv v \pm 1 \pmod{L}$ . In order to consider a parallel process on  $G_n$ , it is also convenient to modify the way in which we count the stages of the process. In the process as previously described, an unhappy pair of nodes of opposite type are selected at each stage, which may then swap (in fact will swap for the values of  $\tau$  considered here). Since we shall now have a situation in which the same node  $u$  may be unhappy in  $C_n$  but happy in  $G_n$ , or vice versa, it becomes convenient to consider a process in which two nodes are selected uniformly at random at each stage, which will only then swap if they are of opposite type and both are unhappy. Of course this makes no real difference to the evolution of the system, except for the way in which we count the stages. The parallel process on  $G_n$  then unfolds as follows: when  $u$  and  $v$  are selected for a potential swap in  $C_n$ , they swap in  $G_n$  if they are both of opposite type and are unhappy in  $G_n$ . Now let us see how the configuration of  $G_n$  at any stage can be described by a  $2^L$ -dimensional vector. We let  $2^L$  denote the set of binary strings of length  $L$ . For each node  $u$  and each stage  $s$  we define a string  $\tau_{u,s} \in 2^L$ :  $\tau_{u,s}(0) = 1$  if  $u$  is of type  $\alpha$  at stage  $s$ , otherwise  $\tau_{u,s}(0) = 0$ , and then  $\tau_{u,s}(1) = 1$  if the node to the right of  $u$  in  $G_n$  is of type  $\alpha$  at stage  $s$ , and so on.

For each  $\sigma \in 2^L$  we define  $\zeta_\sigma(s)$  to be the number of nodes  $u$  such that  $\tau_{u,s} = \sigma$ . Then the  $2^L$ -dimensional vector  $\zeta(s)$  in which the components are the values  $\zeta_\sigma(s)$  for  $\sigma \in 2^L$  describes the configuration at stage  $s$  up to isomorphism.

Our first task now is to show that for large  $L$  and large  $n$  compared to  $L$ , the processes on  $G_n$  and  $C_n$  do not diverge too quickly. To this end we inductively define a set of *tainted* nodes for each stage  $s$ , denoted  $T(s)$ . These are nodes whose neighbourhoods might possibly look different in  $G_n$  and  $C_n$ .  $T(0)$  is the set of  $u$  such that  $-w \leq u \bmod L < w$  (we assume that the swapping process begins at stage  $s = 1$ ). If  $u$  and  $v$  are chosen for a potential swap at stage  $s > 0$  and are both untainted, then  $T(s) = T(s-1)$ . Otherwise  $T(s)$  is the union of  $T(s-1)$  with the set of all nodes which are in the neighbourhood of  $u$  or  $v$  in either  $G_n$  or  $C_n$ . Immediately it is clear that  $|T(s) - T(s-1)| < 4(2w+1)$ . In fact, by counting more carefully those nodes which belong to the neighbourhoods of  $u$  and  $v$  in *both* graphs, we get  $|T(s) - T(s-1)| < 2(3w+1)$ , so that, since  $w$  is large:

$$(4.28.1) \quad |T(s) - T(s-1)| < 7w.$$

The next lemma provides precisely the kind of probabilistic bound on the number of tainted nodes at any given stage which we will need later.

**Lemma 4.29.** *The following conditions hold at every stage  $s \geq 0$ :*

- (1) *The expected number of tainted nodes at stage  $s$  is bounded above by  $e^{14ws/n} \left(\frac{2w-1}{L}\right) n$ .*
- (2) *The probability that  $\frac{|T(s)|}{n} > 2e^{14ws/n} \left(\frac{2w-1}{L}\right)$  is at most  $e^{-wn/(21L^2)}$ .*

**Proof.** Let  $\rho(s) = |T(s)|$ . In order to prove (1) first, let  $u$  and  $v$  be the nodes chosen for a swap at stage  $s+1$ . Let  $\gamma$  be the probability that either of these nodes are tainted (at the end of stage  $s$ ). The probability that  $u$  is tainted is  $\rho(s)/n$ , and similarly for  $v$ , so  $\gamma \leq 2\rho(s)/n$ . Now if neither of  $u$  or  $v$  is tainted then  $\rho(s+1) = \rho(s)$ , and otherwise, by (4.28.1),  $\rho(s+1) < \rho(s) + 7w$ . We therefore have:

$$\mathbf{E}[\rho(s+1) \mid \rho(s)] < (1 - \gamma)\rho(s) + \gamma(\rho(s) + 7w) \leq \rho(s) + 2\frac{\rho(s)}{n}7w = \rho(s) \cdot \left(1 + \frac{14w}{n}\right).$$

For  $x > 0$ ,  $1 + x < e^x$ , giving:

$$\mathbf{E}[\rho(s+1) \mid \rho(s)] < \rho(s) \cdot e^{14w/n}.$$

So far, then, we have established that the sequence of random variables  $Y_s = \rho(s)e^{-14ws/n}$  is a supermartingale. This suffices to give (1) since  $Y_0 = \left(\frac{2w-1}{L}\right) n$ .

In order to establish (2) we make use of the following fact, as proved in [21] for example. If  $X_1, \dots, X_s$  is a supermartingale with  $X_0 = 0$  and  $|X_i - X_{i-1}| \leq c_i$  for all  $i$  and constants  $c_i$ , then for all  $\Omega > 0$ ,

$$(4.29.1) \quad \mathbf{P}(X_s \geq \Omega) \leq \exp\left(-\frac{\Omega^2}{2 \sum_{i=1}^s c_i^2}\right).$$

If we put  $X_s = \frac{Y_s}{n} - \frac{2w-1}{L}$  then the sequence  $X_0, X_1, \dots$  is a supermartingale and  $X_0 = 0$ . In order to apply (4.29.1) we need a bound on the values  $|X_i - X_{i-1}|$ . To this end we consider the sum:

$$|Y_s - e^{-14w/n}Y_s| + |Y_{s+1} - e^{-14w/n}Y_s|.$$

Now

$$Y_s - e^{-14w/n}Y_s = Y_s(1 - e^{-14w/n}) < Y_s 14w/n,$$

since for  $x > 0$ ,  $1 - e^{-x} < x$ . Also,

$$Y_{s+1} - e^{-14w/n}Y_s = e^{-14(s+1)w/n}(\rho(s+1) - \rho(s)) \leq 7we^{-14(s+1)w/n},$$

by (4.28.1). Thus,

$$|Y_s - e^{-14w/n}Y_s| + |Y_{s+1} - e^{-14w/n}Y_s| < 14wY_s/n + 7we^{-14w(s+1)/n} < 21we^{-14ws/n}.$$

We may therefore put  $c_s = \frac{21}{n}we^{-14ws/n}$ . Now since for  $0 < x < 1$ ,  $\frac{1}{1-x} = 1 + x + x^2 + x^3 \dots$  we conclude that  $\sum_{s=0}^{\infty} c_s^2 = (21w/n)^2(1 - e^{-28w/n})^{-1} < 21w/n$ , the latter inequality following from the fact that for small positive  $x$ ,  $1 - e^{-x} > 3x/4$ .

Finally, applying (4.29.1), we get:

$$\begin{aligned} \mathbf{P}\left(\frac{\rho(s)}{n} > 2e^{14ws/n} \left(\frac{2w-1}{L}\right)\right) &= \mathbf{P}\left(\frac{Y_s e^{14ws/n}}{n} > 2e^{14ws/n} \left(\frac{2w-1}{L}\right)\right) \\ &= \mathbf{P}\left(X_s + \left(\frac{2w-1}{L}\right) > 2 \left(\frac{2w-1}{L}\right)\right) \\ &= \mathbf{P}\left(X_s > \left(\frac{2w-1}{L}\right)\right) \\ &\leq \exp\left(-\frac{n(2w-1)^2}{42wL^2}\right) \\ &< \exp\left(-\frac{nw}{21L^2}\right). \end{aligned}$$

This establishes (2) as required.  $\square$

So Lemma 4.29 tells us that for fixed  $x$ , we can ensure with high probability that the proportion of nodes which are tainted by stage  $xn$  is very small, so long as we take  $L$  large and  $n$  large compared to  $L$ . With this in place, we now concentrate for a while on the

processes on the graphs  $G_n$ . In order to approximate these processes by the solution to a system of differential equations, we first of all have to draw up reasonable candidates for the differential equations to be used. To this end we begin by considering the conditional expectations  $\mathbf{E}[\zeta_\sigma(s+1) \mid \zeta(s)]$ . One can form an expression for this expectation, simply by listing all of the different ways in which the number of nodes  $u$  with  $\tau_u = \sigma$  can increase or decrease at a given stage and establishing their probabilities (we formally defined  $\tau_{u,s}$  rather than  $\tau_u$ , but shall omit  $s$  where no ambiguity results). Consider a proposed swap between  $u$  and  $v$  which belong to different cycles, such that  $\tau_u = \sigma'$  and  $\tau_v = \sigma''$ . Define  $a(\sigma, \sigma', \sigma'')$  to be the net change resulting from this proposed swap, in the number of nodes  $u'$  for which  $\tau_{u'} = \sigma$ . Note that the net change will be zero unless  $u$  and  $v$  are of opposite type and are both unhappy, and that in any case  $|a(\sigma, \sigma', \sigma'')| < 2L$ . Then:

$$(4.29.2) \quad \mathbf{E}[\zeta_\sigma(s+1) - \zeta_\sigma(s) \mid \zeta(s)] = \sum_{\sigma', \sigma''} a(\sigma, \sigma', \sigma'') \frac{\zeta_{\sigma'}(s) \zeta_{\sigma''}(s)}{n^2} + O\left(\frac{L^2}{n}\right).$$

The  $O\left(\frac{L^2}{n}\right)$  correction term accounts for the possibility that the nodes selected for a potential swap might belong to the same cycle, the probability of this occurring being  $L/n$  and the net change then being bounded by  $L$ . We next perform some scaling, so that the variables reach fixed functions in the limit as  $n \rightarrow \infty$ : for each  $\sigma$  we want a real function  $z_\sigma(s)$  to model the behaviour of  $\frac{1}{n}\zeta_\sigma(sn)$ . Now (4.29.2) suggests the following system of differential equations for the functions  $z_\sigma$ :

$$(4.29.3) \quad z'_\sigma(s) = \sum_{\sigma', \sigma''} a(\sigma, \sigma', \sigma'') z_{\sigma'}(s) z_{\sigma''}(s).$$

$$(4.29.4) \quad z_\sigma(0) = \frac{1}{n} \zeta_\sigma(0).$$

Note that for any solution to (4.29.3) and (4.29.4), if we put  $y(s) = \sum_\sigma z_\sigma(s)$ , then:

$$y'(s) = \sum_\sigma z'_\sigma(s) = \sum_{\sigma', \sigma''} \sum_\sigma a(\sigma, \sigma', \sigma'') z_{\sigma'}(s) z_{\sigma''}(s) = \sum_{\sigma', \sigma''} 0 = 0.$$

Thus, since  $y(0) = 1$ , we also have  $y(s) = 1$  for all  $s$  in the domain. This in turn means that :

$$(4.29.5) \quad |z'_\sigma(s)| = \left| \sum_{\sigma'} \sum_{\sigma''} a(\sigma, \sigma', \sigma'') z_{\sigma'}(s) z_{\sigma''}(s) \right| \leq \left| \sum_{\sigma'} \sum_{\sigma''} 2L z_{\sigma'}(s) z_{\sigma''}(s) \right| = \left| \sum_{\sigma'} 2L z_{\sigma'}(s) \right| = 2L$$

Suppose for now that we are interested in the first  $xn$  steps of the process on  $G_n$ . Later we shall choose a value for  $x$  which suits our needs. For now though, this means that we are interested in solutions to (4.29.3) and (4.29.4) for  $s \in [0, x]$ . Let  $\Pi$  be the set of

points  $(s, z_1, \dots, z_{2^L})$  such that  $s \in [0, x]$  and, for each  $i \leq 2^L$ ,  $|z_i| \leq 2Lx + 1$ . Let  $D$  be a bounded connected open subset of  $\mathbb{R}^{2^L+1}$  containing  $\Pi$  and such that the minimum  $\ell^\infty$  distance between the boundary of  $D$  and any point in  $\Pi$  is bounded below (and where the  $\ell^\infty$  distance between two vectors is the maximum difference between corresponding components). For the remainder of this section, let all the strings in  $2^L$  be enumerated as  $\sigma_1, \sigma_2, \dots, \sigma_{2^L}$ . Then by standard results in the theory of differential equations (see [14] for example), the fact that each of the functions

$$f_i(s, z_1, \dots, z_{2^L}) := \sum_{j,k} a(\sigma_i, \sigma_j, \sigma_k) z_j z_k$$

satisfies a Lipschitz condition for each argument  $z_r$  in  $D$ , together with (4.29.5), means that there exists a unique set of functions  $\{z_{\sigma_i}(s) \mid \sigma_i \in 2^L\}$  defined on the interval  $[0, x]$  and which satisfy (4.29.3) and (4.29.4) for all  $s \in [0, x]$ .

The following lemma is just a slightly stripped down version of Theorem 5.1 from [21].

**Lemma 4.30.** *Let  $D$  and  $x$  be as above. Suppose that for all  $n$ ,  $\lambda(n) > L^2/n$  and that as  $n \rightarrow \infty$ ,  $\lambda(n) \rightarrow 0$ . Consider the unique set of functions  $\{z_\sigma \mid \sigma \in 2^L\}$  defined on the interval  $[0, x]$  which satisfy (4.29.4) and (4.29.3) for all  $s \in [0, x]$ . Then the following holds with probability  $1 - O(\frac{2^{L+1}L}{\lambda(n)} \exp(-\frac{n\lambda(n)^3}{8L^3}))$ :*

$$\zeta_\sigma(s) = nz_\sigma(s/n) + O(n\lambda(n))$$

uniformly for  $0 \leq s \leq xn$ .

Actually, in the statement Theorem 5.1 in [21], the conclusion is only guaranteed for those  $s$  for which  $(s, z_{\sigma_1}(s), \dots, z_{\sigma_{2^L}}(s))$  is within  $\ell^\infty$  distance  $C\lambda(n)$  of the boundary of  $D$  for some sufficiently large constant  $C$  (where  $C$  is independent from  $n$ ). In our case, however, (4.29.5) together with the way in which we defined  $D$  removes these complications. Since we required that the minimum  $\ell^\infty$  distance between  $D$  and  $\Pi$  be bounded below, and since  $\lambda(n) \rightarrow 0$  as  $n \rightarrow \infty$ , the required condition is automatically satisfied for all sufficiently large  $n$  and all  $s \in [0, x]$  (and then the constant in the  $O(\frac{2^{L+1}L}{\lambda(n)} \exp(-\frac{n\lambda(n)^3}{8L^3}))$  term can be modified to ensure the theorem holds for all  $n$ ).

Consider now the symmetry properties which must be satisfied by the functions  $z_\sigma$ . For each  $\sigma \in 2^L$  let  $\bar{\sigma}$  be the string which results from changing every 1 in  $\sigma$  to 0, and changing every 0 to 1. For  $\mathbf{z} \in \mathbb{R}^{2^L}$ , let  $\iota(\mathbf{z})$  be the vector which results from swapping the  $i$ th component with the  $j$ th component whenever  $\bar{\sigma}_i = \sigma_j$  (recalling our enumeration of the strings of length  $2^L$  from before). If we have  $\iota(\zeta(0)) = \zeta(0)$  (which we probably will not, but for large  $n$  we shall have something close to it with high probability), then the symmetry properties of (4.29.3) guarantee that for all  $s \in [0, x]$ ,  $\iota(\zeta(s)) = \zeta(s)$ . Define  $u(\sigma) = 1$  if any node  $u$  such that  $\tau_u = \sigma$  is an unhappy  $\alpha$  node,  $u(\sigma) = -1$  if any node  $u$  such that  $\tau_u = \sigma$  is an unhappy  $\beta$  node, and  $u(\sigma) = 0$  otherwise. For  $\mathbf{z} = (z_1, \dots, z_{2^L}) \in \mathbb{R}^{2^L}$  consider the linear function:



$$\Delta(\mathbf{z}) = \sum_i u(\sigma_i) z_i.$$

Then  $\Delta(\zeta(s)/n)$  is the difference at stage  $s$ , between the fraction of nodes which are unhappy and of type  $\alpha$ , and the fraction which are unhappy and of type  $\beta$ . Let  $\mathbf{z}(s)$  be the vector solution to (4.29.3) and (4.29.4), i.e. with  $i$ th component  $z_{\sigma_i}(s)$ . If  $\iota(\zeta(0)) = \zeta(0)$  then we have that  $\Delta(\mathbf{z}(s)) = 0$  for all  $s \in [0, x]$ .

In the above we considered the case that  $\iota(\zeta(0)) = \zeta(0)$ . We now have to deal with the fact that this will probably not hold exactly. Recall that we are interested in the processes on  $C_n$  and  $G_n$  up stage  $xn$  for some fixed  $x$ , which we are yet to specify, but which will not depend on  $n$  or  $L$ . For now suppose that we want to ensure the following modified version of our ultimate aim ( $\diamond$ ) (one of the differences being that this modified statement refers to a difference in numbers rather than a ratio):

- ( $\dagger$ ) For fixed  $x > 0$  and any  $\epsilon_1 > 0$  the following holds for the process on  $C_n$ , with probability  $> 1 - \epsilon_1$  for all sufficiently large  $n$ : at all stages  $\leq xn$  the difference between the fraction of nodes which are unhappy and of type  $\alpha$  and the fraction which are unhappy and of type  $\beta$ , is less than  $\epsilon_1$ .

To establish ( $\dagger$ ) take  $\epsilon_2 < \epsilon_1$ . It suffices to ensure that with probability  $> 1 - \epsilon_2$ ,  $|\Delta(\mathbf{z}(s))| < \epsilon_2$  for all  $s \in [0, x]$ , since then we can take  $L$  large and  $n$  sufficiently large compared to  $L$  and apply Lemmas 4.29 and 4.30 (putting  $\lambda(n) = \max\{\frac{L^2+1}{n}, n^{-1/4}\}$ , for example) to get the required result for  $C_n$ .

Let  $S$  denote the set of vectors  $\mathbf{y} = (y_1, \dots, y_{2L})$  such that the solution to (4.29.3) with  $\mathbf{z}(0) = \mathbf{y}$ , satisfies  $|\Delta(\mathbf{z}(s))| < \epsilon_2$  for all  $s \in [0, x]$ . Let  $\Sigma = \{\mathbf{y} \mid \iota(\mathbf{y}) = \mathbf{y}, \forall i, y_i \geq 0, \sum_i y_i = 1\}$ . By the continuity properties of ordinary differential equations (see for example [3]) it follows that  $S$  is an open set, so since  $\Sigma$  is compact there exists  $d_0$  such that every point within distance  $d_0$  of  $\Sigma$  belongs to  $S$ . The it follows directly from the law of large numbers that when  $n$  is sufficiently large compared to  $L$ ,  $\zeta(0)/n$  is within distance  $d_0$  of  $\Sigma$  with probability  $> 1 - \epsilon_2$ . This establishes ( $\dagger$ ) as required.

Finally, we have to specify the value  $x$ . Recall that our aim is to prove ( $\diamond$ ), as specified previously. Let  $l_1$  be the length of the interval  $[l_{k_0}, r_{k_0}]$ . Note that (for fixed  $w$  and  $\epsilon$ , where  $\epsilon$  is the value for which we are proving Theorem 1.2, and which plays a role in the value of  $k_0$ , and for  $\epsilon'$  as in ( $\diamond$ )), we can take  $l_2$  such that for all sufficiently large  $n$ , the probability that either  $|l_{k_0} - u_0| > l_2/2$  or  $|u_0 - r_{k_0}| > l_2/2$  is  $\ll \epsilon'$ . Taking  $l_3 \gg l_2/\epsilon'$ , consider now the first stage  $s_1$  at which, for some  $\gamma \in \{\alpha, \beta\}$ , the number of unhappy  $\gamma$  nodes as a proportion of  $n$  is less than  $1/l_3$ . Taking  $\epsilon'' \ll \epsilon'$  and putting  $x > (2w+1)(l_3)^2/\epsilon''$ , suppose towards a contradiction that  $\mathbf{P}(s_1 > xn) > \epsilon''$ . In that case, with probability  $> \epsilon''$ , the following holds at each stage  $1 \leq s \leq xn$ : given the configuration at the end of stage  $s-1$ , the probability at stage  $s$  of choosing two nodes which do swap is at least  $\frac{1}{(l_3)^2}$ . The expected number of stages  $\leq xn$  at which swaps do occur is therefore  $> (2w+1)n$ , which

gives the required contradiction since, according to the observation made in Section 2, there can be at most  $(2w + 1)n$  stages at which a swap occurs. Applying  $(\dagger)$  to  $\epsilon_1 \ll \epsilon'/l_3$  then establishes that with probability approaching 1 as  $n \rightarrow \infty$ ,  $s_1 < xn$  and at all stages  $\leq xn$  the difference between the fraction of nodes which are unhappy and of type  $\alpha$  and the fraction which are unhappy and of type  $\beta$ , is less than  $\epsilon_1$ . When the latter conditions hold, this means that at all stages up to  $s_1$  the ratio between the number of unhappy  $\alpha$  nodes and the number of unhappy  $\beta$  nodes lies in the interval  $[1 - \epsilon', 1 + \epsilon']$ . Now choosing  $l_3 \gg l_2/\epsilon'$  as we did previously, means that, since  $u_0$  was chosen uniformly at random, and given that the total proportion of nodes which are unhappy at stage  $s_1$  is at most  $\frac{2+\epsilon'}{l_3}$ , the probability that there are any unhappy nodes in the interval  $[u_0 - \lceil l_2/2 \rceil, u_0 + \lceil l_2/2 \rceil]$  (or in the almost certainly smaller interval  $[l_{k_0}, r_{k_0}]$ ) at stage  $s_1$  is  $\ll \epsilon'$ . Thus we have established  $(\diamond)$ , as required.

### 5. THE CASE $\tau = \frac{1}{2}$

This case was dealt with in [2], for both the simple and standard models. Although we have talked about threshold behaviour occurring at both  $\tau = \kappa$  and  $\tau = \frac{1}{2}$ , it is worth noting that these two thresholds are of essentially different types. For a fixed large  $w$ , as one increases the value of  $\tau$  past the lower threshold  $\kappa$ , one will very suddenly observe entirely different results for the final configuration – from a situation in which almost no type changing occurs, one moves to a situation where  $u$  chosen uniformly at random can be expected to belong to a firewall of length exponential in  $w$  in the final configuration. As one increases  $\tau$  gradually up to  $\frac{1}{2}$ , however, what one observes is a smooth decrease in the expected length of the firewall to which  $u$  belongs, with the situation only reversing very suddenly when  $\tau > \frac{1}{2}$ . In order to observe the difference in behaviour for when  $\tau$  is just less than  $\frac{1}{2}$  or when  $\tau = \frac{1}{2}$ , one should consider instead these two fixed  $\tau$ , and gradually increase  $w$ . Then, when  $\tau < \frac{1}{2}$ , the expected length of the firewall to which  $u$  belongs grows exponentially, while for  $\tau = \frac{1}{2}$  it grows only polynomially in  $w$ . Figure 6 displays this difference in behaviour, and also illustrates the distinct mechanisms involved for the two cases. These diagrams illustrate the process exactly as in Figure 1, the difference being that now  $n = 1000000$  rather than 50000. It is also interesting to observe the difference in the way that the firewall spreads as a ‘wave’ emanating from unhappy nodes in the initial configuration for different  $\tau < \frac{1}{2}$ . When  $\tau$  is just above  $\kappa$  the wave is sharply defined, (as illustrated in Figure 1), much less so when  $\tau$  is just below  $\frac{1}{2}$ .

### 6. THE CASE $\tau > \frac{1}{2}$

We consider first the standard model. Note that if  $\frac{1}{2} < \tau \leq \frac{w+1}{2w+1}$  then the process is identical to that for  $\tau = \frac{1}{2}$ . We therefore assume in what follows that  $w$  is sufficiently large to ensure  $\tau > \frac{w+1}{2w+1}$ , which is equivalent to the condition that adjacent nodes of opposite types cannot both be happy. We show that with probability  $1 - \epsilon$  the starting configuration is such that complete segregation results with probability 1, where  $\epsilon \rightarrow 0$  as

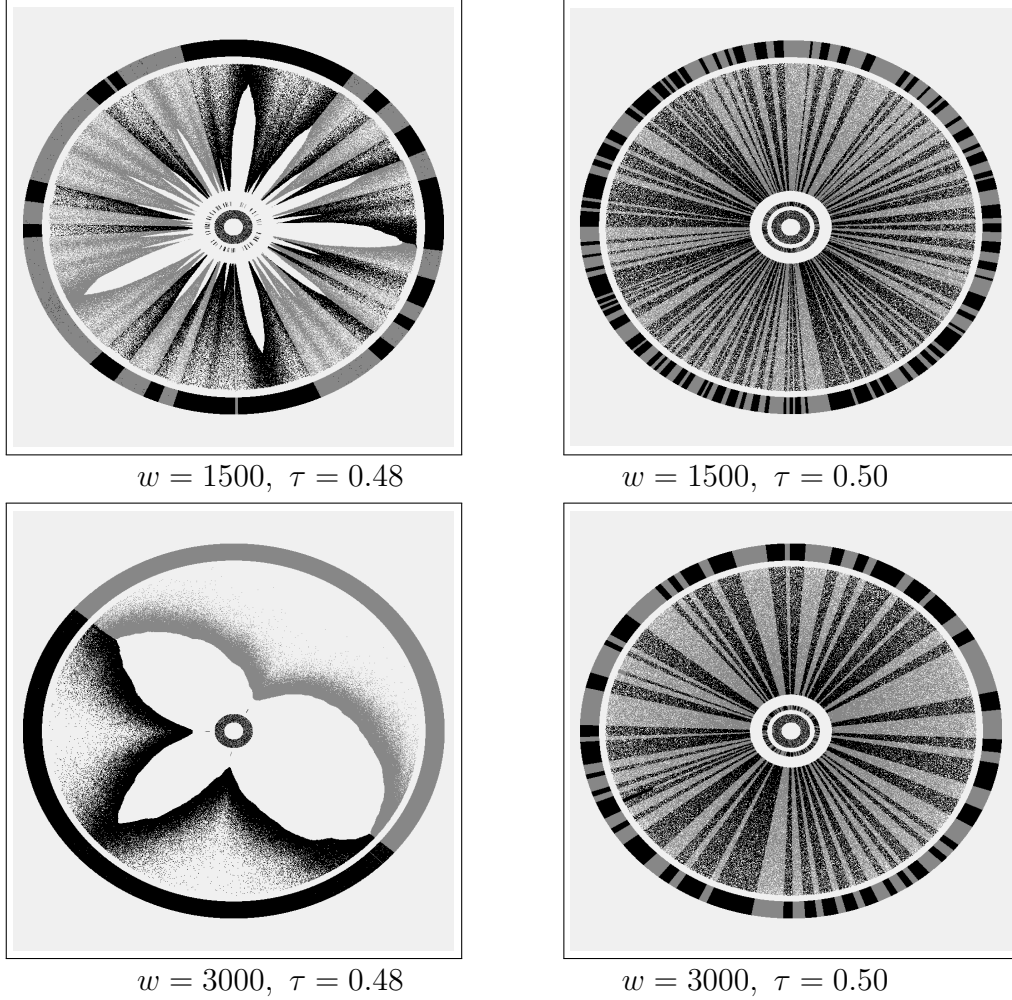


FIGURE 6. The second kind of threshold behaviour.

$n \rightarrow \infty$ . Recall that complete segregation refers to any configuration in which all  $\alpha$  nodes belong to a single run, and that, as observed in Section 2, once a completely segregated configuration is reached all future configurations must be completely segregated.

Let  $x$  be the number of  $\alpha$  nodes in the initial configuration, and let  $x_p = x/n$ . Our task is to show that for sufficiently large  $n$  it is possible to reach a completely segregated configuration from any other, so long as  $x_p$  is close to  $\frac{1}{2}$ . To prove this, however, one must be able to ensure the existence of unhappy individuals of both types at each step along the way. To this end we first of all prove that, with probability tending to 1 as  $n \rightarrow \infty$ ,  $x$  satisfies the property that there are unhappy  $\alpha$  nodes in *any* configuration on a ring of size  $n$  with  $x$  many  $\alpha$  nodes. While this may sound intuitively rather obvious, it (perhaps surprisingly) takes a little work to prove.

**Proving that (when  $x_p$  is close to  $\frac{1}{2}$ ) there are unhappy  $\alpha$  nodes in any configuration.** Rather than considering the nodes to be arranged in a circle, it is temporarily useful to suppose instead that they are arranged in a line which extends infinitely far to the right, with positions indexed by natural numbers (starting with 0). So we consider a fixed configuration in which each natural number is either of type  $\alpha$  or  $\beta$ . It is convenient to suppose further that the node at  $w$  is of type  $\alpha$ , and that all  $\alpha$  nodes  $\geq w$  are happy. From these assumptions, we will deduce the following bound on the proportion of  $\alpha$  nodes:

(†) For any  $\epsilon > 0$  there exists  $\ell$  such that for all  $\ell' \geq \ell$ , the proportion of  $\alpha$  nodes in the interval  $[0, \ell']$  is  $\geq \frac{w+1}{2w} - \epsilon$ .

This suffices because the existence of any circle in which all  $\alpha$  nodes are happy and in which  $0 < x_p < \frac{w+1}{2w}$  (the latter condition on  $x_p$  being satisfied almost certainly for large enough  $n$ , by the law of large numbers) would give a contradiction, since this circle could then be cut and infinitely many copies placed end to end, giving a counter example to (†).

In order to prove (†), it will be useful to consider the right and left parts of any neighbourhood  $\mathcal{N}(u)$  for  $u \geq w$ . We define  $\mathcal{L}(u)$  (respectively  $\mathcal{R}(u)$ ) to be the leftmost (rightmost)  $w$ -many nodes in  $\mathcal{N}(u)$ . Note that any  $\alpha$  node  $u \geq w$  must have  $\alpha$  nodes in both  $\mathcal{L}(u)$  and  $\mathcal{R}(u)$ . Also, any  $\beta$  nodes  $v > w$  must have  $\alpha$  nodes in  $\mathcal{L}(v)$  and in  $\mathcal{R}(v)$ , otherwise the next  $\alpha$  node to the left or right (respectively) would not be happy.

We define a sequence of  $\alpha$  nodes. Let  $u_0 = w$ . Given  $u_k$  define  $u_{k+1}$  to be the rightmost  $\alpha$  node in  $\mathcal{N}_k := \mathcal{N}(u_k)$ . Also define  $\mathcal{L}_k := \mathcal{L}(u_k)$  and  $\mathcal{R}_k = \mathcal{R}(u_k)$ . Now, for  $m \geq 1$  define  $I_m := \bigcup_{k=0}^m \mathcal{N}_k$  and  $S_m := \sum_{k=0}^m \Theta(\mathcal{N}_k)$ . Since each  $u_k$  is happy,  $\Theta(\mathcal{N}_k) \geq 3$ , giving  $S_m \geq 3(m+1)$ . This doesn't immediately tell us the bias on the interval  $I_m$ , however, because nodes may have been counted multiple times in forming the sum  $S_m$ . We therefore want to consider the way in which the neighbourhoods  $\mathcal{N}_k$  overlap. For  $k \geq 1$ , define  $\mathcal{N}'_k := \mathcal{R}_{k-1} \cap \mathcal{L}_{k+1}$  ( $= \mathcal{N}_{k-1} \cap \mathcal{N}_{k+1}$ ). Notice that  $u_k \in \mathcal{N}'_k$ , but that  $u_{k-1} \notin \mathcal{N}'_k$  and  $u_{k+1} \notin \mathcal{N}'_k$ . We similarly partition each  $\mathcal{N}'_k$  into a left and right part. Define  $\mathcal{L}'_k := \mathcal{N}'_k \cap \mathcal{L}_k$  and  $\mathcal{R}'_k := \mathcal{N}'_k \cap \mathcal{R}_k$ . It is immediate that if  $\mathcal{R}'_k$  is non-empty then all nodes in it are of type  $\beta$ . Now we partition the nodes in  $I_m$  according to the number of times they are counted in forming the sum  $S_m$ , i.e. the number of  $\mathcal{N}_k$  that they belong to for  $k \leq m$ . We define  $J_r^m$  to be the set of  $u \in I_m$  which belong to exactly  $r$  distinct neighbourhoods  $\mathcal{N}_k$  for  $k \leq m$ . Notice that  $\bigcup_{r \geq 3} J_r^m \subseteq \bigcup_{k=1}^{m-1} \mathcal{N}'_k$ . On the other hand,  $u_k \in \bigcup_{r \geq 3} J_r^m$  for  $1 \leq k \leq m-1$ . We now want to assess the size and bias of each of the various  $J_r^m$ .

$J_1^m$ . We have  $J_1^m \subseteq \mathcal{L}(u_0) \cup \mathcal{R}(u_m)$ , so  $|J_1^m| \leq 2w$ .

$J_4^m$ . We show that only  $\beta$  nodes can belong to  $J_4^m$ . Suppose that  $u \in J_4^m$  is of type  $\alpha$ . Then  $u \in \mathcal{N}'_k$  for some  $k$  with  $1 \leq k \leq m-1$ , and by the remark above  $u \notin \mathcal{R}'_k$ . So  $u = u_k$  or  $u \in \mathcal{L}'_k$ . In either case,  $u \in \mathcal{N}_{k-2}$  will contradict the definition of  $u_{k-1}$ , while  $u \in \mathcal{N}_{k+2}$  will contradict the definition of  $u_{k+1}$ , giving the required contradiction.

$J_r^m$ ,  $r \geq 5$ . Suppose  $u \in \mathcal{N}_{k-2} \cap \mathcal{N}_{k-1} \cap \mathcal{N}_k \cap \mathcal{N}_{k+1} \cap \mathcal{N}_{k+2}$ . Then either  $u_k \in \mathcal{N}_{k-2}$  contradicting the definition of  $u_{k-1}$  or  $u_k \in \mathcal{N}_{k+2}$  contradicting that of  $u_{k+1}$ . Thus  $J_r^m$  is empty for  $r \geq 5$ .

$J_3^m$ . We claim that  $\Theta(J_3^m) < m + w$ . From our analysis of  $J_r^m$  for  $r > 3$  it follows that we must have  $u_k \in J_3^m$  for  $1 \leq k \leq m-1$ . So  $\Theta(J_3^m) \leq m-1 + \sum_{k=1}^{m-1} |\mathcal{L}'_k| - \sum_{k=1}^{m-1} |\mathcal{R}'_k|$ . It therefore suffices to show that  $\sum_{k=1}^{m-1} (|\mathcal{L}'_k| - |\mathcal{R}'_k|)$  is bounded by  $w$ . For  $k \geq 1$  define  $d_k := u_k - u_{k-1}$ . Then  $|\mathcal{R}'_k| = w - d_k$  and  $|\mathcal{L}'_k| = w - d_{k+1}$ . Hence  $|\mathcal{L}'_k| = |\mathcal{R}'_{k+1}|$ , and  $\sum_{k=1}^{m-1} (|\mathcal{L}'_k| - |\mathcal{R}'_k|) = |\mathcal{L}'_{m-1}| - |\mathcal{R}'_1| \leq w$ , proving the claim.

Putting these facts together, we get that  $2\Theta(I_m) > 3m - 2w - m - w = 2m - 3w$ , so  $\Theta(I_m) > m - \frac{3}{2}w$ . For  $m > \frac{3}{2}w$  choose a partition,  $I_m = \Pi_0 \cup \Pi_1 \cup \Pi_2$ , so that  $|\Pi_0| = |\Pi_1|$ ,  $\Pi_0$  contains only  $\beta$  nodes, and  $\Pi_1$  and  $\Pi_2$  contain only  $\alpha$  nodes (meaning that  $|\Pi_2| > m - \frac{3}{2}w$ ). Then since  $|I_m| \leq 2w + 1 + mw$ , the proportion of the nodes in  $I_m$  which belong to  $\Pi_2$  is greater than

$$\frac{m - \frac{3}{2}w}{2w + 1 + mw} > \frac{m}{m+3} \cdot \frac{1}{w} - \frac{3}{2m+6}$$

which tends to  $\frac{1}{w}$  as  $m \rightarrow \infty$ . Thus, for each  $\epsilon > 0$ , there exists  $m$  such that for all  $m' \geq m$ , the proportion of  $\alpha$  nodes in  $I_{m'}$  is  $\geq \frac{w+1}{2w} - \epsilon$ , giving  $(\dagger)$ , as required.

Actually, it is convenient to be able to have a little more than the existence of unhappy nodes of both types to work with. In the analysis that follows it will be convenient to work with configurations in which we are guaranteed the existence of unhappy nodes outside any interval of length  $4w + 1$ . We can get this at the expense of assuming  $n$  to be reasonably large compared to  $w$ . Given any circle configuration in which all  $\alpha$  nodes outside an interval of length  $4w + 1$  are happy, cut the circle at the left end of the interval of length  $4w + 1$  and consider the individuals in the circle to lie in a line with sites indexed by natural numbers  $< n$  (so those in the interval of length  $4w + 1$  occupy  $[0, 4w]$ ). Performing an almost identical analysis we still conclude that, for any  $\epsilon > 0$ , so long as  $n$  is sufficiently large, there exists  $\ell$  such that for all  $\ell' \geq \ell$  (with  $\ell' < n$ ), the proportion of  $\alpha$  nodes in the interval  $[0, \ell']$  is  $\geq \frac{w+1}{2w} - \epsilon$ . We conclude that with probability  $1 - \epsilon$  the initial configuration on a circle of size  $n$  will have a value  $x$  sufficient to ensure unhappy nodes of both types outside any interval of length  $4w + 1$  in *any* configuration with the same number of  $\alpha$  and  $\beta$  nodes, where  $\epsilon \rightarrow 0$  as  $n \rightarrow \infty$ .

**Moving to complete segregation.** So now we return to considering nodes arranged in a circle, and from now on we assume that all configurations considered have unhappy nodes of both types outside any given interval of length  $4w + 1$ . We build a list of configurations from which it is possible to reach complete segregation. Consider first any configuration which is not completely segregated, but which has a run of length at least  $2w$ . Without loss of generality, suppose that this is a run of  $\alpha$  nodes occupying the interval  $[a, b]$ , where this interval is chosen to be of maximum possible length. If the nodes  $a$  and  $b$  are both happy then the length of the interval ensures that all nodes in the run are happy – this

follows by induction on the distance from the edge of the interval. In this case let  $u$  be an unhappy  $\alpha$  node and let  $c \in \{a, b\}$  be distance at least  $w + 1$  from  $u$ . Then  $u$  and the  $\beta$  neighbour of  $c$  may legally be swapped, increasing the length of the run by at least 1. So suppose instead that at least one of the individuals  $a$  and  $b$  is not happy, and without loss of generality suppose that  $a$  has bias less than or equal to  $b$ . Then  $a$  and  $b + 1$  may legally be swapped. Performing this swap causes position  $b + 1$  to have at least the same bias as  $b$  did before the swap, and causes  $a + 1$  to have at most the same bias as  $a$  did before the swap. Thus, the swap has the effect of shifting the run one position to the right and may be repeated until the length of the run is increased by at least 1, i.e. for successive  $i \geq 0$  we can swap the nodes  $a + i$  and  $b + i + 1$ , so long as the latter is of type  $\beta$ . The first stage at which the latter is of type  $\alpha$  the length of the run has been increased. Putting these observations together, we conclude that from any configuration which has a run of length at least  $2w$  it is possible to reach full segregation.

Next consider a configuration in which the longest run  $[a, b]$  is of length at least  $w$ , but strictly less than  $2w$ . We shall suppose that  $[a, b]$  contains  $\alpha$  nodes, the case for  $\beta$  nodes is similar. Let  $c$  be the first  $\alpha$  node strictly to the left of  $a$ . If  $c$  is unhappy, then we may legally swap  $c$  and  $a - 1$ , strictly increasing the length of the longest run. If  $c$  is happy then the distance between  $c$  and  $a$  is at most  $w$  and we may successively swap unhappy  $\alpha$  nodes from outside the interval  $[c - w, a + 2w]$  with the nodes  $c + i$  for  $1 \leq i < a - c$  (starting with  $i = 1$  and proceeding in order), in order to strictly increase the length of the longest run. This follows because as each node  $c + i$  performs the swap, it will become happy.

It remains to show that we can always move to a configuration with a run of length at least  $w$ . Given any configuration (satisfying the condition that any other configuration reached by swapping unhappy nodes has unhappy nodes of both types outside any given interval of length  $4w + 1$ ), we first of all perform a procedure which selects a number of individuals with relatively unfavourable bias. Let  $u_0$  be an  $\alpha$  node with least possible bias. Given  $u_k$  for  $k < w^2$  choose  $u_{k+1}$  outside  $\bigcup_{j \leq k} \mathcal{N}(u_j)$  which has least possible bias amongst all the  $\alpha$  nodes at such sites (we can suppose that  $n$  is large enough that such a choice is possible). Once  $u_k$  is defined for each  $k \leq w^2$ , choose  $v_0$  outside  $\bigcup_{j \leq w^2} \mathcal{N}(u_j)$ , which has the greatest possible bias amongst all  $\beta$  nodes at such sites. Given  $v_k$  for  $k < w^2$  choose  $v_{k+1}$  outside  $\bigcup_{j \leq w^2} \mathcal{N}(u_j)$  and outside  $\bigcup_{j \leq k} \mathcal{N}(v_j)$  which has greatest possible bias amongst all the  $\beta$  nodes at such sites. Now choose an interval  $[a, b]$  of length  $w$  such that  $[a - w, b + w]$  has no intersection with any of the neighbourhoods  $\mathcal{N}(u_j)$  or  $\mathcal{N}(v_j)$  for  $j \leq w^2$  (again assuming  $n$  is sufficiently large).

The point of this procedure was to provide a pool of individuals which we can use for legal swaps. Now we perform another iteration, which produces a run in the interval  $[a, b]$ . At any point during the iteration we say that a neighbourhood  $\mathcal{N}(u_j)$  or  $\mathcal{N}(v_j)$  is *tarnished* if any node in that neighbourhood has been involved in a swap since the iteration began. Using the notation from Definition 3.2, if  $\gamma = \beta$  then  $\gamma^* = \alpha$ , and otherwise  $\gamma^* = \beta$ .



**Step 0.** Let  $\gamma = \alpha$  if  $a$  is of type  $\alpha$ , otherwise let  $\gamma = \beta$ . At any given point in the iteration the value of  $\gamma$  specifies the type of run that we are looking to produce, which may change as the iteration progresses. Set  $i := 0$ .

**Step  $s > 0$ .** We are given a configuration in which all nodes in  $[a, a + i]$  are of type  $\gamma$ , and no nodes in  $[a - w, b + w]$  have yet been involved in swaps, except possibly those in  $[a, a + i]$ . Also, at most  $i^2$  of the neighbourhoods  $\mathcal{N}(u_k)$  or  $\mathcal{N}(v_k)$  are tarnished. Let  $j$  be the least  $> i$  such that  $a + j$  is not of type  $\gamma$ . If  $j \geq w$  then the iteration is complete, and we carry out no further instructions. Otherwise, if  $a + j$  is unhappy, we divide into two subcases. If  $a + j - 1$  is happy then we can legally swap any unhappy  $\gamma$  node from outside the interval  $[a - w, b + w]$  with  $a + j$ . If  $a + j - 1$  is unhappy, we can select any  $u_k$  or  $v_k$  of type  $\gamma$  whose neighbourhood is not tarnished and legally swap that node with  $a + j$ . In order to see why this is the case suppose that  $\gamma = \alpha$ , the case for  $\beta$  is similar. Firstly, the fact that at most  $i^2$  of these neighbourhoods are tarnished means that we certainly have untarnished neighbourhoods of the appropriate type to choose from. Also, before the iteration began any  $u_k$  had at most bias  $\Theta(\mathcal{N}(a + j - 1))$  (as that value was defined then). At the present point in the iteration,  $\mathcal{N}(u_k)$  being untarnished means that the neighbourhood  $\mathcal{N}(u_k)$  is unchanged, while any node in the interval  $[a - w, b + w]$  which is now of different type than before, is presently of type  $\alpha$ . Swapping  $u_k$  with  $a + j$  will cause it to have bias at least the same as that which  $a + j - 1$  had before the swap. In either of these two subcases, once the swap is performed, redefine  $i := j$  and proceed to step  $s + 1$ . The final case to consider is that the individual at  $a + j$  is happy. In this case, if  $\gamma = \alpha$  then redefine  $\gamma = \beta$ , or if instead  $\gamma = \beta$  then redefine  $\gamma = \alpha$ . For successive values of  $k$ ,  $1 \leq k \leq j$ , swap an unhappy  $\gamma$  individual from outside the interval  $[a - w, b + w]$  with  $a + j - k$ , causing that node to become happy. Once this sequence of swaps is complete, redefine  $i := j$  and proceed to step  $s + 1$ .

**The simple model.** We wish to show that, whatever the initial configuration, with probability 1 a configuration in which all nodes are of the same type is reached. For the purposes of this discussion, for  $\gamma \in \{\alpha, \beta\}$ , we shall say that  $\gamma$  is a *minority type* if there are at most as many nodes of type  $\gamma$  as of type  $\gamma^*$  (with  $\gamma^*$  defined as in Section 6). The argument in Section 6 actually suffices to show that if  $\gamma$  is a minority type, then there exists at least one  $\gamma$  node which is unhappy, and which would have at least as many neighbours of its own (new) type if it changed type. Given any configuration, one can then select  $\gamma$  of minority type and successively select nodes of type  $\gamma$  which can legally have their type swapped, until all nodes are of type  $\gamma^*$ .

## REFERENCES

- [1] I. Benenson, E. Hatna, and E. Or, From Schelling to spatially explicit modelling of urban ethnic and economic residential dynamics, *Sociological Methods & Research*, 37(4):463, 2009.
- [2] C. Brandt, N. Immorlica, G. Kamath, R. Kleinberg, An Analysis of One-Dimensional Schelling Segregation, *STOC 2012*.
- [3] F. Brauer and J. A. Nohel, The Qualitative Theory of Ordinary Differential Equations. Dover, 1989.



- [4] Bruch, E. E. and R. D. Mare, Neighborhood Choice and Neighborhood Change, *American Journal of Sociology*, 112, 667-709, 2006.
- [5] W. Clark, Residential segregation in American cities: A review and interpretation, *Population Research and Policy Review*, 5(2):95-127, 1986.
- [6] M. Emerson, K. Chai, and G. Yancey, Does race matter in residential segregation? Exploring the preferences of white americans, *American Sociological Review*, pages 922-935, 2001.
- [7] Epstein, J. M. and R. Axtell, Growing Artificial Societies: Social Science from the Bottom Up. Washington, DC: Brookings Institution Press, 1996.
- [8] Fossett, M., Ethnic Preferences, Social Distance Dynamics, and Residential Segregation: Theoretical Explorations Using Simulation Analysis, *Journal of Mathematical Sociology*, 30, 185-273, 2006
- [9] Fagiolo, G., M. Valente, and N. J. Vriend, Segregation in Networks, *Journal of Economic Behaviour and Organization*, 64, 316-336, 2007.
- [10] R. Farley, C. Steeh, T. Jackson, M. Krysan, and K. Reeves, Continued racial residential segregation in Detroit: Chocolate city, vanilla suburbs revisited, *Journal of Housing Research*, 4(1):1-38, 1993.
- [11] S. Gerhold, L. Glebsky, C. Schneider, H. Weiss, and B. Zimmermann, Limit states for one- dimensional Schelling segregation models, *Communications in Nonlinear Science and Numerical Simulation*, 13(10):2236-2245, 2008.
- [12] S. Grauwin, F. Goffette-Nagot, and P. Jensen, Dynamic models of residential segregation: Brief review, analytical resolution and study of the introduction of coordination. Arxiv preprint arXiv:0907.1777, 2009.
- [13] A. D. Henry, P. Pralat, and C. Zhang, Emergence of segregation in evolving social networks. *Proceedings of the National Academy of Sciences*, 108(21):8605-8610, May 2011.
- [14] W. Hurewicz, Lectures on Ordinary Differential Equations, M.I.T. Press, Cambridge Massachusetts, 1958.
- [15] B. D. McKay, On Littlewood's estimate for the binomial distribution, *Adv. Appl. Prob.*, 21 (1989) 475-478. Available at <http://cs.anu.edu.au/~bdm/papers/littlewood2.pdf>.
- [16] R. Pancs and N. Vriend, Schellings spatial proximity model of segregation revisited, *Journal of Public Economics*, 91(1-2):1-24, 2007.
- [17] T. Schelling, Models of segregation, *The American Economic Review*, pages 488-493, 1969.
- [18] A.M.Turing, The Chemical Basis of Morphogenesis , *Phil. Trans. R. Soc. London B*, 237, pp 37-72, 1952.
- [19] M. White, Segregation and diversity measures in population distribution, *Population index*, 52(2):198-221, 1986.
- [20] D. Vinković and A. Kirman, A Physical Analogue of the Schelling Model, *Proceedings of the National Academy of Sciences*, no. 51, volume 103, 19261-19265, 2006.
- [21] N. C. Wormald, Differential equations for random processes and random graphs, *Annals of Applied Probability*, 5:1217-1235, 1995.
- [22] Young, H.P., Individual Strategy and Social Structure: An Evolutionary Theory of Institutions. Princeton, NJ: Princeton University Press, 1998.
- [23] J. Zhang, A dynamic model of residential segregation, *Journal of Mathematical Sociology*, 28(3):147-170, 2004.
- [24] J. Zhang, Residential segregation in an all-integrationist world, *Journal of Economic Behavior & Organization*, 54(4):533-550, 2004.
- [25] J. Zhang, Tipping and residential segregation: A unified Schelling model, *Journal of Regional Science*, 51:167-193, Feb. 2011.

**George Barmpalias**, STATE KEY LAB OF COMPUTER SCIENCE, INSTITUTE OF SOFTWARE, CHINESE ACADEMY OF SCIENCES, BEIJING 100190, P.O. BOX 8718, PEOPLE'S REPUBLIC OF CHINA.

*E-mail address:* `barmpalias@gmail.com`

*URL:* <http://barmpalias.net>

**Richard Elwes**, SCHOOL OF MATHEMATICS, UNIVERSITY OF LEEDS, LS2 9JT LEEDS, UNITED KINGDOM.

*E-mail address:* `r.h.elwes@leeds.ac.uk`

**Andy Lewis-Pye**, SCHOOL OF MATHEMATICS, UNIVERSITY OF LEEDS, LS2 9JT LEEDS, UNITED KINGDOM.

*E-mail address:* `andy@aemlewis.com`

*URL:* <http://aemlewis.co.uk>